J. Mol. Biol. (1994) **236**, 759–785

High Resolution Structures of Holo and Apo Formate Dehydrogenase

Victor S. Lamzin[1,2], Zbigniew Dauter[1], Vladimir O. Popov[2] Emil H. Harutyunyan[3] and Keith S. Wilson[1]

[1]*European Molecular Biology Laboratory (EMBL) c/o DESY, Notkestraße 85, 22603 Hamburg, Germany*

[2]*Institute of Biochemistry, Russian Academy of Sciences Leninsky pr. 33, Moscow 117071, Russia*

[3]*Institute of Crystallography, Russian Academy of Sciences Leninsky pr. 59, Moscow 117333, Russia*

Three-dimensional crystal structures of holo (ternary complex enzyme-NAD-azide) and apo NAD-dependent dimeric formate dehydrogenase (FDH) from the methylotrophic bacterium *Pseudomonas* sp. 101 have been refined to *R* factors of 11·7% and 14·8% at 2·05 and 1·80 Å resolution, respectively. The estimated root-mean-square error in atomic co-ordinates is 0·11 Å for holo and 0·18 Å for apo. X-ray data were collected from single crystals using an imaging plate scanner and synchrotron radiation. In both crystal forms there is a dimer in the asymmetric unit. Both structures show essentially 2-fold molecular symmetry.

NAD binding causes movement of the catalytic domain and ordering of the C terminus, where a new helix appears. This completes formation of the enzyme active centre in holo FDH. NAD is bound in the cleft separating the domains and mainly interacts with residues from the co-enzyme binding domain. In apo FDH these residues are held in essentially the same conformation by water molecules occupying the NAD binding region. An azide molecule is located near the point of catalysis, the C4 atom of the nicotinamide moiety of NAD, and overlaps with the proposed formate binding site. There is an extensive channel running from the active site to the protein surface and this is supposed to be used by substrate to reach the active centre after NAD has already bound. The structure of the active site and a hypothetical catalytic mechanism are discussed. Sequence homology of FDH with other NAD-dependent formate dehydrogenases and some D-specific dehydrogenases is discussed on the basis of the FDH three-dimensional structure.

Keywords: formate dehydrogenase; NAD-dependent dehydrogenases, crystal structure; domain movement; active centre

1. Introduction

NAD-dependent formate dehydrogenase (FDH†, EC 1.2.1.2) from the methylotrophic bacterium *Pseudomonas* sp. 101 catalyses the oxidation of formate anion with concomitant reduction of NAD to NADH. The catalytic reaction appears to involve cleavage of a single carbon–hydrogen bond in the formate anion and the formation of a single new one in the co-enzyme. Thus, FDH catalyses the simplest NAD-dependent dehydrogenation of carbonyl compounds. Kinetic experiments suggest that no water molecules or ions participate in the catalysis (Egorov *et al.*, 1979). FDH is a dimer with two chemically identical subunits each possessing an independent active centre (Popov *et al.*, 1978). Each subunit has 393 residues according to protein sequence determination (Popov *et al.*, 1990) or 400 from the sequence of the FDH gene, where seven extra residues have been found at the C terminus (Tishkov *et al.*, 1991).

X-ray studies of FDH aim to provide a structural basis for the catalytic mechanism. Recently we reported the crystal structure of hFDH (ternary complex of enzyme with NAD and azide inhibitor) at 3·0 Å resolution (Lamzin *et al.*, 1992). X-ray data at higher resolution are now available both for hFDH (2·05 Å) and for aFDH (1·8 Å). Crystal structure determination of the two forms at high resolu-

† Abbreviations used: FDH, NAD-dependent bacterial formate dehydrogenase (EC 1.2.1.2); aFDH, apo FDH; hFDH, holo FDH (ternary complex enzyme-NAD-azide); ARP, automated refinement procedure; SA, simulated annealing; r.m.s., root-mean-square.

0022–2836/94/080759–27 $08.00/0

tion is important for several reasons. To date only a few crystal structures of NAD-dependent dehydrogenases in the Protein Data Bank (Bernstein *et al.*, 1977) have been analysed at a resolution of 2·0 Å or better. They are glyceraldehyde-3-phosphate dehydrogenase (co-ordinate set 1GD1, Skarzynski *et al.* (1987) at 1·8 Å resolution) and lactate dehydrogenase (6LDH: Abad-Zapatero *et al.* (1987); 1LLD: Iwata & Otha (1993); 9LDT: Dunn *et al.* (1991) at 2·0 Å). There are additional representatives of dinucleotide binding enzymes such as dihydrofolate reductase (3DFR and 4DFR: Bolin *et al.* (1982) at 1·7 Å; 1DRA, 1DRB, 3DRC: Bystroff & Kraut (1991) at 1·9 Å and 1DRF: Oefner *et al.* (1988) at 2·0 Å), cholesterol oxidase (1COX: Vrielink *et al.* (1991) at 1·8 Å) and glutathione reductase (3GRS: Karplus & Schulz (1987) at 1·5 Å). Other structures of NAD(P)/FAD-dependent enzymes such as alcohol dehydrogenase, ferredoxin reductase, isocitrate dehydrogenase, malate dehydrogenase and hydroxybenzoate hydroxylase and their complexes have only been analysed at medium resolution, where location of atoms participating in the active site cannot be determined precisely.

FDH differs substantially from other NAD-dependent dehydrogenases (such as alcohol, glyceraldehyde-3-phosphate, lactate and malate dehydrogenases) in quaternary and domain structure and has very low primary structure homology to these proteins (Lamzin *et al.*, 1992). Recently, FDH was found to have a strong sequence homology with several so-called D-specific 2-hydroxy acid dehydrogenases (Vinals *et al.*, 1993). At present the three-dimensional structure of these D-enzymes is unknown. They catalyse reactions with D-hydroxy acids, which involve both hydride and proton transfer, as L-specific enzymes. In contrast, the reaction catalysed by FDH is not stereochemically specific towards the substrate and comprises hydride transfer only. Knowledge of the FDH structure will provide information about the overall fold and active centre of D-specific enzymes in comparison with L-specific dehydrogenases. Comparative studies of aFDH (pre-transition state of the reaction) and hFDH (analogue of the transition state) will provide detailed information for analysis of mechanism.

Here we describe the crystallization, data collection and refinement at high resolution for both aFDH and hFDH. Two alternative molecular replacement and refinement protocols of aFDH were carried out and a comparison of the results is given. The secondary structure of FDH was described previously (Lamzin *et al.*, 1992) and further details revealed at high resolution as well as analysis of hydrogen bond pattern are now presented. Conformational changes both in the enzyme active centre and the whole fold are considered in relation to NAD binding. Detailed analysis of the FDH active centre, proposed substrate binding site and important features of catalysis are discussed.

2. Crystallization

FDH was isolated from the methylotrophic bacterium *Pseudomonas* sp. 101 according to the purification procedure published earlier (Lamzin *et al.*, 1992). EDTA was used as a stabilizer of the enzyme. The protein was at least 98% pure according to polyacrylamide gel electrophoresis.

Crystals of hFDH and aFDH diffracting to high resolution were grown according to the protocol described earlier. In brief, crystallization was carried out at room temperature using vapour diffusion (Davies & Segal, 1971) with ammonium sulphate as a precipitant and 5% by volume of 2-methyl-2,4-pentanediol. For both hFDH and aFDH the starting volume in the batch crystallization was about 0·5 ml. Crystals of hFDH were grown in the presence of 5 mM NAD and 5 mM sodium azide and reached dimensions of 3·0 mm × 0·6 mm × 0·4 mm after four weeks. The final ammonium sulphate concentration was 50% saturation. The space group is $P2_12_12_1$ with cell dimensions $a = 116{\cdot}0$ Å, $b = 113{\cdot}3$ Å and $c = 63{\cdot}4$ Å. Crystals of aFDH grew to dimensions of up to 1·2 mm × 0·6 mm × 0·3 mm after three weeks, with a final concentration of ammonium sulphate of 45 % saturation. The space group is $P2_1$ with cell dimensions $a = 110{\cdot}5$ Å, $b = 54{\cdot}7$ Å, $c = 70{\cdot}3$ Å and $\beta = 101{\cdot}9°$. Taking the molecular mass of the FDH dimer as 86·3 kDa (Popov *et al.*, 1990) the packing density is 2·4 Å^3/Da for both apo and holo forms. In both forms the asymmetric unit contains a dimer.

3. Data Collection and Processing

High resolution data for both hFDH and aFDH were collected in two (low and high resolution) data

Table 1
Summary of data collection

	hFDH		aFDH	
	High resolution	Low resolution	High resolution	Low resolution
Maximum resolution (Å)	2·05	2·9	1·8	2·5
Crystal-plate distance (mm)	200	300	150	260
Oscillation per image (°)	1·0	1·5	1·0	2·0
Number of images	90	60	180	90
Crystal mount	b*	b*	b*	b*

Table 2
Summary of the data processing

	hFDH	aFDH
R_{merge} overall (%)	4·5	5·8
R_{merge} in highest resolution shell (%)	7·5	12·9
Number of reflections:		
Total measured	206,376	308,786
Fully recorded	173,305	246,853
Partially recorded	33,071	61,933
Unique	51,142	73,882
Completeness (all reflections):		
Overall (%)	96·5	97·1
In highest resolution shell (%)	96·3	94·4
Completeness (only reflections with $I_o > 3\sigma$):		
Overall (%)	90·3	89·7
In highest resolution shell (%)	84·7	69·3
Temperature factor from Wilson (1942) plot (Å^2)	11·1	13·9

The highest resolution shell for hFDH is from 2·10 to 2·05 Å, for aFDH from 1·85 to 1·80 Å. R_{merge} is defined as $\Sigma|I-\langle I\rangle|/\Sigma I$, where I is the measured intensity, $\langle I\rangle$ is the average intensity for equivalent reflections and the summation is over equivalent measurements.

sets from single crystals. An imaging plate scanner on the X31 synchrotron beam line at the EMBL outstation on the DORIS storage ring, DESY, Hamburg, was used with radiation of wavelength 1·009 Å. A summary of the data collection is presented in Table 1.

The data were processed using the MOSFLM package modified for image plates (Leslie, 1992). Data merging and all subsequent calculations used the CCP4 programme suite (CCP4, 1979) unless stated otherwise. The data processing is summarized in Table 2. R_{merge} as a function of resolution is shown in Figure 1. R_{merge} is actually higher for aFDH data compared to hFDH, especially at low resolution. This fact is discussed below in terms of higher thermal motion in the crystal of aFDH.

4. Methods

(a) *Refinement*

Simulated annealing (SA) refinement was performed with X-PLOR (Brünger *et al.*, 1987) and was used at step 1 in the refinement of hFDH and at Step 4 in the restrained refinement of aFDH. In all stages of SA refinement the charges on the glutamate, aspartate, lysine and arginine side-chains were turned off. Stereochemically restrained least-squares minimization was carried out with PROLSQ (Konnert & Hendrickson, 1980). A fast Fourier transform for calculation of structure factors and gradients was used as originally suggested by Agarwal (1978) and implemented by Baker & Dodson (1980). In all steps all X-ray data were used with no cutoff on amplitude. A rigid-body refinement was performed with CORELS (Sussman *et al.*, 1977). An Automated Refinement Procedure (ARP), which is a combination of least-squares refinement with automatic updating of the model on the basis of Fourier syntheses (Lamzin & Wilson, 1993) was applied as an alternative refinement of aFDH and the protocol is described below.

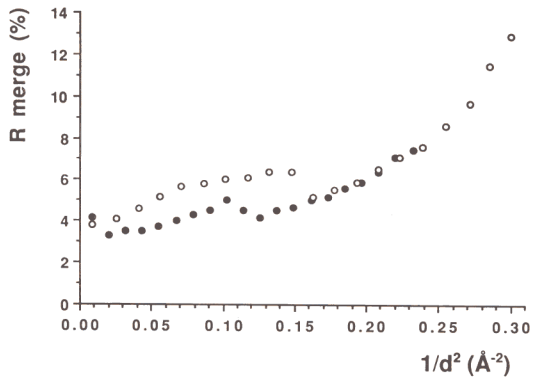


Figure 1. Merging R factor as a function of resolution. filled circles, hFDH; open circles, aFDH.

(b) *Density averaging*

Iterative averaging of the electron density was used at steps 7 and 8 in the refinement of hFDH. The ($3F_o-2F_c$, α_c) density calculated at 2·05 Å resolution was averaged according to the non-crystallographic two-fold symmetry. The density outside spheres with 2 Å radius around each atomic centre was set to the average value expected for solvent. The resulting density was transformed to the asymmetric unit and used to calculate new phases. This was iterated several times. The procedure applied is based on density averaging as suggested by Bricogne (1974) but the co-ordinates of the current model were used as the molecular boundary.

(c) *Correction of the model*

Rebuilding of the model was performed with FRODO (Jones, 1978). Removal and addition of solvent molecules was done automatically (Lamzin & Wilson, 1993). Removal was carried out on the basis of density values and distance criteria (minimum H-bond distance and non-bonded contacts were set to 2·2 Å, maximum H-bond distance of 3·3 Å was used). New solvent molecules were selected from the difference Fourier synthesis taking into account the value of the electron density, H-bond distances and non-bonded contacts. The lowest level of difference density for addition of new water molecules in the last steps of the refinement was 3σ for both aFDH and hFDH models. The Ramachandran (ϕ/ψ) plot was

computed using the PROCHECK software (Morris *et al.*, 1992).

5. Refinement of Holo FDH

The starting model for refinement at high resolution was hFDH built into the 3·0 Å MIR map and refined to an R factor of 27·1% without solvent (Lamzin *et al.*, 1992). The course of the refinement is presented in Table 3.

The final model of hFDH comprises 6013 protein atoms, two NAD and two azide molecules, 788 waters and one sulphate ion. There are 391 residues in the A and 383 residues in the B subunit visible in the electron density. All residues are in a single conformation and all atoms have been assigned unit occupancy. Pro312 and Pro314 are in the *cis* conformation.

6. Refinement of Apo FDH

(a) *Molecular replacement*

The aFDH structure was solved using molecular replacement with the hFDH model after step 15 of refinement (Table 3). Only atoms of Ala1 to Lys383

Table 3

Course of the refinement of holo FDH at 2·05 Å resolution

Step	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
No. of cycles	1	2–5	6–8	9–11	12–17	18–23	24–28	29–33	34–36	37–39	40–42	43–45	46–48	49–51	52–53	54–55
Manual intervention before this step	No	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes
No. of water molecules	—	51	371	313	313	281	389	471	628	663	756	790	788	788	788	788
No. of atoms	6176	6227	6547	6489	6489	6455	6239	6513	6858	6755	6863	6902	6900	6990	6900	6900
R factor (%)	28·8	27·4	26·8	26·4	25·7	24·5	20·3	16·4	15·1	13·9	13·0	12·5	12·4	11·9	11·8	11·7

Step 1, SA was performed in the resolution range 8·0 to 2·05 Å. There were 6088 protein atoms corresponding to 2×393 residues and 88 atoms of 2 NAD molecules in the model. All atomic temperature factors were reset to 15 Å^2 and hydrogen atoms were generated and included. The initial R factor for this model was 37·2%. 160 cycles of energy minimization were run (the "prepare" stage). The R factor slightly increased to 38·9%. This probably indicates the presence of energetically unfavoured regions in the model originating from medium resolution refinement. The "heat" stage was carried out at 3000 K in 4000 steps with 1 fs dynamics at each step. The "cool" stage was performed as a "fast cooling" in 20 steps with 50 fs dynamics when the effective temperature decreased from 3000 K to 300 K. The "final" stage included 40 cycles of energy minimization similar to the "prepare" stage. This resulted in an R factor of 33·1%. Individual temperature factors were refined and 40 cycles of energy minimization were run. The resulting R factor was 28·8%.

Steps 2 and 3, Hereafter the model was refined with stereochemical restraints using PROLSQ. The lowest resolution limit was extended to 10 Å. All hydrogen atoms were excluded. Only solvent molecules were updated.

Step 4, The model was inspected for the first time using FRODO. In both subunits residues Val138, Thr139 and Val145 were replaced by Thr, Val and Cys, respectively, as suggested from the gene sequence of FDH. Side-chain conformations for about 25 residues were adjusted; 68 water molecules were rejected and 10 new ones were added. The X-ray weight was increased from 1·0 to 1·5.

Step 5, The conformations of 47 side-chains were adjusted.

Step 6, Only solvent molecules were changed with removal of 115 and addition of 83 waters.

Step 7, Because the model still contained errors which could not be corrected from the density maps, an iterative non-crystallographic averaging of the $(3F_o-2F_c, \alpha_c)$ electron density at 2.05 Å resolution was performed, as described above. The resulting density showed substantial improvement and major rebuilding was carried out. An extensive region of chain from Glu64 to Pro79 was shifted forwards by 2 residues in both subunits. More than 30 residues were rebuilt and in several places the main-chain carbonyl group was flipped. In both subunits residues Ser77, His215 and Val216 were replaced by Asp, Val and His, respectively, according to the gene sequence of FDH. Density corresponding to the azide inhibitor was found in both active centres and these two molecules were included. The correct conformation of several residues at the C terminus was still not clearly visible, so residues Val368 to Ala393 in both subunits were omitted. 89 water molecules were rejected and 197 new ones included.

Step 8, Electron density averaging was again applied. Inspection of the resulting density showed some of the C-terminal residues. Val368 to Ser382 were rebuilt in both subunits. The rest of the C-terminal residues could not be traced with confidence. Conformations of side-chains for 61 residues were adjusted. Two proline residues, Pro312 and Pro314, were found to be in the *cis* conformation in both subunits, 83 water molecules were rejected and 165 new ones were found.

Step 9, Residues Lys383 to Ser390 were built in both subunits, 32 side-chains were slightly corrected, 78 water molecules were rejected and 239 new ones found.

Step 10, Several residues at the C terminus were omitted: Ser390 in the A subunit and Gly384 to Ser390 in the B subunit. About 60 side-chains were adjusted. Several atoms in NAD and azide were moved to higher density, 79 water molecules were rejected and 114 new ones were found.

Step 11, Only minor corrections were made. Several side-chain conformations, mostly for residues located at the surface of the protein, were adjusted. Residues Ser390 and Glu391 in the A subunit were reintroduced, 34 water molecules were removed and 127 new ones were found. The X-ray weight was reduced to 1·0 to improve the geometry.

Step 12 and 13, Only solvent molecules were updated. The R factor fell to 12·4%. At this stage the restrained refinement of hFDH was deemed to have converged. However, an attempt to improve the model by including hydrogen atoms has been made and is described in the following steps.

Step 14, Before each cycle all hydrogen atoms were generated in idealized positions and their partial contribution was included in the restrained refinement of non-hydrogen atoms. The R factor fell to 11·9% with some improvement in the geometry, especially chiral volumes.

Step 15, Only solvent was updated. The model was again refined taking fixed hydrogen contributions into account. The R factor changed only to 11·8% with no improvement of geometry, suggesting the refinement had converged. Steps 14 and 15 of the refinement with hydrogen atoms resulted in an improvement in R factor of 0·6%.

Step 16, After the aFDH has been refined (see below) the following minor changes were introduced in the holo model. The amide groups of Asn66 and Gln96 were rotated by 180° in both subunits resulting in better packing and H-bonds. Asp327 was changed to asparagine as suggested from the gene sequence. The R factor fell to 11·7%.

Table 4
First steps of the crystallographic refinement of aFDH

Molecular replacement using:	Whole molecule			Individual domains		
Step	1	2	3	1	2	3
Programme	CORELS	PROLSQ	PROLSQ	CORELS	PROLSQ	PROLSQ
No of cycles	3	4	7	3	4	7
No. of parameters	36	–	–	36	–	–
No. of atoms	–	5966	5966	–	5966	5966
No. of reflections used	6144	6150	15,557	6144	6150	15,557
Minimum resolution (Å)	8·0	8·0	8·0	8·0	8·0	8·0
Maximum resolution (Å)	4·0	4·0	3·0	4·0	4·0	3·0
R factor at start (%)	47·0	39·5	33·0	40·8	33·6	28·7
R factor at end (%)	38·9	23·3	27·6	32·9	18·9	23·8
Mean deviation from ideality 1-2 distances (Å)	0·018	0·066	0·046	0·018	0·057	0·040

in both subunits were used. NAD, azide and all solvent molecules were excluded. The rotation and translation functions had clear solutions and the latter agreed with one of the possible solutions deduced using a packing function. About 2200 strongest reflections out of 6151 in the resolution range from 8·0 to 4·0 Å were used. Refinement included use of CORELS with rigid bodies being assigned to fragments from residues 1 to 146 and from 334 to 383 (catalytic domain) and from 147 to 333 (co-enzyme binding domain) for each of the subunits followed by further refinement with PROLSQ. However, the results were not very encouraging (left part of Table 4). This was thought to arise most probably from different relative conformations of the domains in the two states of FDH giving problems for molecular replacement.

NAD-dependent dehydrogenases, as well as proteins of some other families, can undergo substantial conformational changes upon ligand binding (e.g. see Eklund & Bränden 1987). For enzymes with clear subdivision into domains, most changes are the result of movement of domains relative to one another as rigid bodies. Earlier experiments with FDH revealed such behaviour (Lamzin *et al.*, 1986). Therefore, we altered the scheme of molecular replacement to obtain independent rotation and translation solutions for the FDH domains.

The use of separate solutions for multi-domain proteins has been reported previously. The structure of the Fab fragment from the auto immune poly(dT)-specific antibody was solved separately for different domains (Cygler & Anderson, 1988). Similarly in the solution of the complex of subtilisin with the inhibitor eglin-C molecular replacement was applied first for protein and then for inhibitor after preliminary refinement of the protein alone and subtraction of its contribution from observed structure factors (Dauter *et al.*, 1991).

Firstly only the two co-enzyme binding domains from the hFDH dimer were used. These are the core of the dimer and comprise approximately half the structure. Solutions for rotation and translation functions were readily found and were approximately the same as when the whole dimer was used. The two co-enzyme binding domains were refined as rigid bodies and then with PROLSQ to an *R* factor of 32·5% at 4·0 Å resolution. The structure factor amplitudes calculated for co-enzyme binding domains were, with proper scaling, subtracted from the observed amplitudes.

Independent solutions now had to be found for the two catalytic domains, each about one quarter of the structure. However, the correct peaks in the rotation function were first for the catalytic domain of the A subunit but only tenth for the catalytic domain of the B subunit. The rotation angles differed somewhat from the previous solutions for co-enzyme binding domains and for the whole molecule. The translation function for "catalytic domain A" was solved easily but not for "catalytic domain B". Thus, only the catalytic domain in the A subunit was added to the model. The relative position along the polar *y* axis was adjusted so as to minimize the deviations from ideality in CA-CA distances at the domain interface, i.e. Asn146 CA to Ser147 CA and Ile333 CA to Ser334 CA in the A subunit were fitted to deviate by not more than 1 Å from expected values. No significant overlaps between catalytic and co-enzyme binding domains were found. The model now included three out of four domains, was refined at 4·0 Å and gave an *R* factor of 24·1%. The structure factors calculated were subtracted from the observed amplitudes.

The rotation function for the remaining domain now had a clear solution corresponding to the highest peak. The translation function was also easily solved. The "extra" translation along the polar axis was adjusted as for the A subunit. Further refinement of aFDH is summarized in Table 4, right, and was much more successful than when molecular replacement was applied to the whole dimer. The r.m.s. deviation in CA positions between the final refined aFDH model and the model after molecular replacement is 0·5 Å for the "domain" molecular replacement solution while it is 1·4 Å for the solution using the whole dimer. The model of aFDH from the "domain" molecular

Table 5

Course of the restrained refinement of aFDH

Step	1	2	3	4	5	6	7	8	9	10	11
No. of cycles	1–3	4–10	11–19	20–20	21–26	27–31	32–34	35–40	41–43	44–49	50–55
Manual intervention before this step	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
No. of water molecules	–	–	–	–	428	656	757	541	595	604	500
No. of atoms	5966	5966	5966	5966	6394	6622	6723	6433	6443	6418	6,314
Maximum resolution (Å)	4·0	3·0	2·4	2·4	2·4	2·0	1·8	1·8	1·8	1·8	1·8
R factor (%)	18·9	23·8	22·0	21·6	16·2	17·9	19·1	18·3	17·4	16·9	16·7

Steps 1, 2 and 3, There were 5971 protein atoms corresponding to 2×383 residues. Only the positional x, y, z parameters for each atom were refined. The resolution range was gradually extended to 2·4 Å and X-ray weighting increased from 0·5 to 1·0.

Step 4, SA refinement was carried out. Taking into account the limited improvement in the refinement of hFDH using X-PLOR at 2.05 Å resolution we now used 2·4 Å as the highest resolution to increase the radius of convergence. Hydrogen atoms were included. All atomic temperature factors were initially reset to 15 Å^2. The initial R factor was 26·2%, 200 cycles of energy minimization were run during the "prepare" stage. The "heat" stage was carried out at 3000 K in 2000 steps with 1 fs dynamics at each step. The "cool" stage was performed as a "slow cooling" in 108 steps with 12·5 fs dynamics when the effective temperature was decreased from 3000 K to 300 K with increment of 25 K. The "final" stage included 40 cycles of energy minimization similar to the "prepare" stage. This resulted in a model with an R factor of 24·9%, only 1·3% lower than for the initial model. Individual temperature factors were refined and the resulting R factor was 21·6% at 2·4 Å resolution.

Steps 5, 6 and 7, In these steps the refinement was carried out using PROLSQ. Only solvent molecules were updated and the resolution was extended to 1·8 Å.

Step 8, 23 side-chain conformations were adjusted. Residues Val368 to Leu373 were assigned zero occupancy because the electron density was not good enough to correct this region, 216 water molecules were rejected.

Step 9, 13 side-chain conformations were corrected. Val368 to Ala372 in the A subunit were rebuilt and Leu373 to Lys383 in the A subunit and Val368 to Ala374 in the B subunit were assigned zero occupancy, 101 water molecules were rejected and 155 new ones added.

Step 10, Major rebuilding of the model was carried out. About 60 residues were corrected including main and side-chain atoms. Residues Leu373 to Lys383 in both subunits were assigned zero occupancy. All residues from Ala1 to Ala372 were present in both chains of the model, 129 water molecules were rejected and 138 new ones added.

Step 11, Side-chain conformations were adjusted for 26 residues, 166 water molecules were rejected and 62 new ones added. The R factor changed slightly to 16·7%. At this stage we decided the restrained refinement was complete.

replacement was used as the starting point for refinement. Two different refinement schemes were used. They are described in detail below.

(b) *Refinement*

The course of restrained refinement of aFDH is shown in Table 5.

Subsequent to the restrained refinement of aFDH, a novel Automatic Refinement Procedure (ARP, see Methods) was developed using these data. Its application is especially powerful when X-ray data are of high quality and extend to a resolution better than 2·0 Å.

Step 1, the initial model was the same as for the restrained refinement above. The ARP protocol has been described (Lamzin & Wilson, 1993). In brief, 24 cycles of ARP were run and all data in the resolution range 10·0 to 1·8 Å were used from the start. Each cycle consisted of one cycle of unrestrained refinement of the x, y, z and B parameters for each atom, followed by removal of the 0·3% of the atoms and addition of 1·0% new atoms on the basis of the density maps. At the end of ARP the R factor fell to 13·8% in the whole resolution range and the ARP model contained 6844 atoms. This number of atoms is actually more than the number of atoms in the model refined with restraints, but the excess seems to be helpful in refinement with ARP.

Initial and ARP models were inspected in the ARP $(3F_o-2F_c, \alpha_c)$ density. Most ARP atoms were located rather close to initial protein atoms. However in several places ARP atoms and the density calculated from these were located 2 to 2·5 Å away from the initial model and showed the correct conformation. The interatomic distances in the ARP model were similar to the corresponding distances in the protein in spite of no restraints being applied. About 90% of the new protein model was automatically constructed using a rebuilding programme (Lamzin & Wilson, 1993). The remaining 10% of the structure consisting of about 600 atoms corresponded to regions with large shifts from the initial model. These were rebuilt manually from the ARP model and $(3F_o-2F_c, \alpha_c)$ density. The protein model included residues Ala1 to Leu373 in both subunits and the total number of atoms was 5830. The other residues at the C terminus were not visible in the ARP map. A total of 514 water molecules were automatically selected. The other 500 ARP atoms were rejected.

The course of post-ARP refinement with geometrical restraints using PROLSQ is presented in Table 6.

After step 3 the results of refinement of aFDH using ARP are now compared with conventional restrained refinement as described in the previous section. In both, the initial model was obtained from the "domain" molecular replacement using the hFDH model followed by refinement at low resolution with each of the four protein domains treated as a rigid body. The conventional refinement took about two months in real time and gave a final model with an R factor of 16·7% and an r.m.s. deviation in bond distances of 0·025 Å. In contrast a

Table 6

Course of the post-ARP refinement of aFDH in the resolution range 10·0 to 1·8 Å

Step	2	3	4	5	6	7	8	9
No. of cycles	1–10	11–14	15–17	20–20	21–23	24–25	26–28	29–34
Manual intervention before this step	Yes	No	Yes	No	Yes	No	No	Yes
No. of water molecules	514	519	811	871	899	881	887	854
No. of atoms	6344	6349	6646	6706	6744	6726	6732	6704
Maximum resolution (Å)	1·8	1·8	1·8	1·8	1·8	1·8	1·8	1·8
R factor (%)	17·6	16·6	15·7	15·4	15·2	14·8	14·7	14·8

Step 2, The resulting protein model, with rather poor geometry (r.m.s. deviation in bond lengths of 0·18 Å), was subjected to restrained refinement with an X-ray weight of 0·7. The initial *R* factor was 23.5% and after 10 cycles fell to 17·6%. The model acquired good geometry.

Step 3, Only solvent molecules were changed. The X-ray weight was increased to 1·0.

Step 4, The conformation for about 20 side-chains was adjusted. Ala374 in the A subunit was included, 20 water molecules were rejected and 312 new solvent molecules found.

Step 5, Only solvent was updated.

Step 6, Conformations of several side-chains were adjusted. Ala374 in the B subunit was built. Two sulphate anions were identified in the electron density, 44 water molecules were rejected and 72 new ones found.

Steps 7 and 8, Only solvent molecules were updated. From these steps the refinement was carried out using contribution from hydrogen atoms as for hFDH.

Step 9, 97 water molecules were removed and 64 new ones added. In both subunits Asp327 was replaced by Asn according to the gene sequence. The X-ray weight was reduced to 0·8 to improve geometry. The refinement resulted in a model with an *R* factor of 14·8%. Here we deemed refinement to be complete.

model with nearly the same characteristics (*R* factor of 16·6%, r.m.s. deviation in bond distances of 0·023 Å) was obtained after only a few days for ARP and rebuilding the 10% of the structure which could not be automatically assigned and to tidy up the stereochemistry with several cycles of restrained refinement. The two models are essentially identical with r.m.s. deviations in positions of CA atoms of 0·09 Å and about 0·19 Å for all protein atoms with temperature factors less than 50 Å^2. The model obtained with ARP was improved further as described in the following steps, Table 6.

The final aFDH model comprises 5840 protein atoms, 854 water molecules and two sulphate ions. There are 384 residues in both A and B subunits visible in the electron density. As in hFDH all residues are in a single conformation and all atoms have unit occupancy. Pro312 and Pro314 are in the *cis* conformation.

7. Description of the Structures

(a) *Accuracy of the models*

The overall characteristics of the models are summarized in Table 7. The *R* factor as a function of resolution for both models is plotted in Figure 2. The increase of *R* factor at low resolution is due to incomplete description of solvent. The behaviour of the *R* factor is nearly the same for both aFDH and hFDH at a resolution of 3·0 Å and lower. At higher resolution the *R* factor for aFDH is markedly higher than for hFDH. This arises from a higher thermal motion of aFDH in the crystal as deduced from the R_{merge} and Wilson plots.

The r.m.s. deviation in bond lengths is 0·016 Å and the largest deviation is 0·102 Å in hFDH. In aFDH the respective values are 0·018 and 0·085 Å. The upper estimate of the r.m.s. error in the co-ordinates derived from σ_A plots (Read, 1986) is 0·11 Å for hFDH and 0·18 Å for aFDH. For hFDH the error in the atomic co-ordinates was further estimated by carrying out ten cycles of unrestrained least-squares refinement of the final model. For this "unrestrained" holo model the r.m.s. deviation in bond lengths from ideality was 0·17 Å. This value divided by $\sqrt{2}$ agrees well with the value obtained from the σ_A plot. The higher estimated error in the atomic co-ordinates for aFDH is a result of its higher thermal motion.

Average atomic temperature factors for the A subunit are shown as a function of residue number in Figure 3. Temperature factors in the B subunit are very similar and the correlation coefficients between average temperature factors per residue in the A and B subunits are 75% for aFDH and 85% for hFDH. The correlation between apo and holo is 73%. The average temperature factors both for main-chain and side-chains are higher almost everywhere for aFDH compared to hFDH. Atoms having temperature factors higher than 50 Å^2 comprise only 6% of the total number of protein atoms in aFDH and less than 3% in hFDH. These corre-

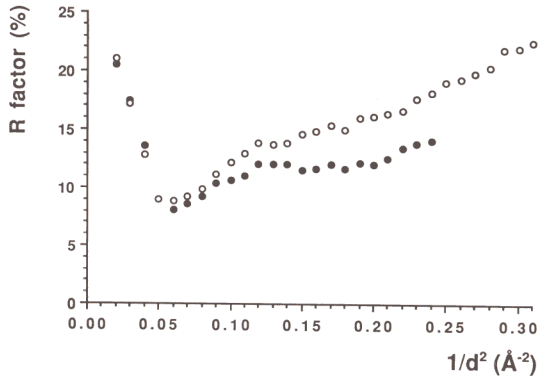


Figure 2. Crystallographic *R* factor as a function of resolution for final models. Filled circles, hFDH; open circles, aFDH.

Table 7
Summary of the refinement and characteristics of final models

		hFDH	aFDH
Resolution range (Å)		10·0–2·05	10·0–1·8
No. of reflections used		50,775	73,551
No. of residues in A and B subunits		391 + 383	374 + 374
No. of atoms		6900	6704
No. of protein atoms		6013	5840
No. of ligand atoms		94	–
No. of water molecules		788	854
No. of sulphate ions		1	2
R factor (%)		11·7	14·8
A. *Mean temperature factor (Å²)*			
Main-chain atoms		10·2	16·4
Side-chain atoms		16·5	24·5
Protein		13·5	20·4
NAD		11·3	–
Azide		18·1	–
Water molecules		38·7	46·5
Sulphate		51·0	47·6
Total		16·4	23·7
B. *Mean deviation from ideality*	Target		
1–2 distances (Å)	0·020	0·016	0·018
1–3 distances (Å)	0·040	0·046	0·043
Planar groups (Å)	0·020	0·016	0·016
Chiral volumes (Å³)	0·150	0·171	0·172
Planar torsion angles (°)	3·0	3·0	3·0
Staggered torsion angles (°)	15·0	15·1	15·3
Orthonormal torsion angles (°)	20·0	32·3	33·0

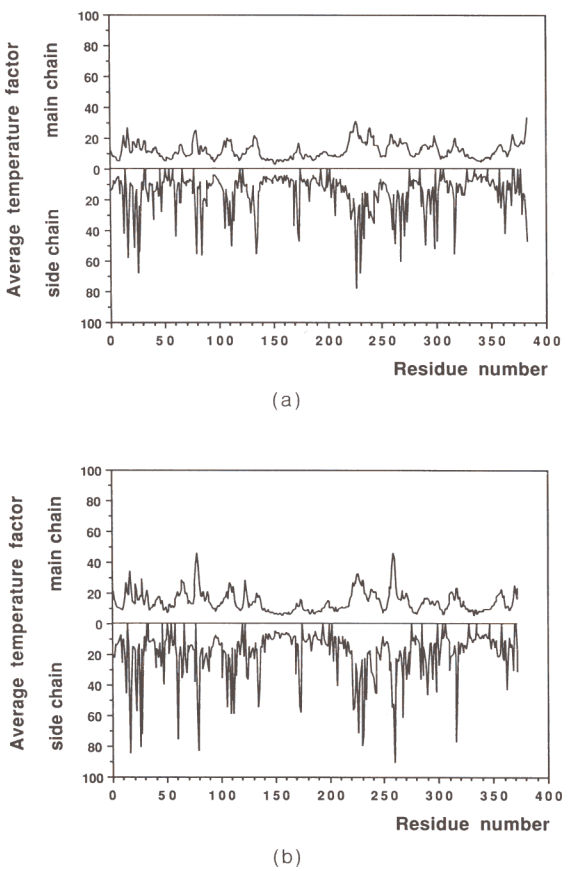


Figure 3. Average temperature factors for main and side-chain atoms for the A subunit of the (a) hFDH and (b) aFDH models.

spond to the poorly ordered residues located at the surface of the protein. Some C-terminal residues in both forms were not visible in the density and not included in the models.

The minimum, maximum and r.m.s. values of difference densities are −0·27, 0·32 and 0·05 e $Å^{-3}$ for hFDH at 2·05 Å resolution. For aFDH these are −0·33, 0·33 and 0·06 e $Å^{-3}$ at 1·8 Å. The correlation coefficient between (F_o, α_c) and (F_c, α_c) maps is 99·0% and 99·2% for aFDH and hFDH, respectively. Histograms of electron density calculated at the atomic positions are presented in Figure 4. The histograms for protein atoms for aFDH and hFDH, look similar except that for aFDH the distribution has a greater mean value and greater dispersion. Histograms of electron densities for water molecules also differ for aFDH and hFDH. Most waters in aFDH have low density and only a small fraction of them have density comparable with protein atoms. For hFDH a relatively larger part of the solvent is as well-defined as protein. This reflects the lower overall temperature factor for hFDH.

FDH primary structures from protein (Popov *et al.*, 1990) and gene (Tishkov *et al.*, 1991) sequencing differ at seven positions, (Table 8). All seven indicated by gene sequencing have been corroborated by the structural data. The side-chains for all seven positions are clearly seen in the electron density. Selection of Asp77, Cys145, Val215 and His216 was easy. The swap of His215-Val216 (protein) to Val215-His216 (gene) was a misprint in the original paper. The choice between Val and Thr (residues 137 and 138) and Asp and Asn (residue 327)

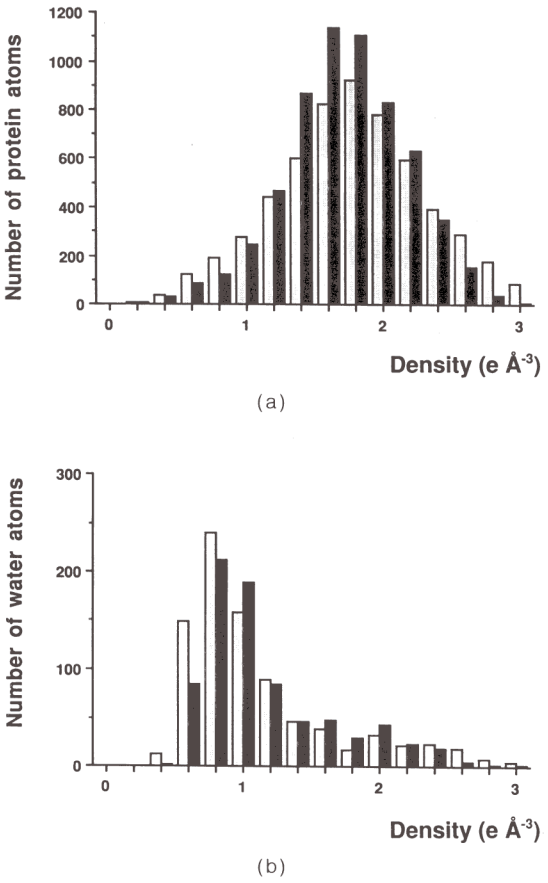


Figure 4. Distribution of the electron densities calculated at the atomic positions. Black columns, hFDH; grey columns, aFDH. (a) Protein atoms; (b) water atoms.

was made from atomic temperature factors in the side-chains. Incorrect choice of nitrogen (Asn/Asp) or carbon (Val/Thr) in place of oxygen and *vice versa* causes considerable inconsistency in the temperature factors. Similarity in these values confirms correctness of the choice made. In addition, Thr138 OG1 forms two H-bonds with solvent molecules. The side-chains of Val137 and Val215 are oriented inside the protein and have no attached water molecules. The appearance of Asp rather than Asn at position 327 in the protein sequence could occur as a result of deamidation in the protein sample. The change of Thr137-Val138 (protein) to Val137-Thr138 (gene) is also easily rationalized as the protein sequence was determined from the C terminus of a rather long peptide, and due to considerable carry over in the course of sequencing the order of the residues may have been misinterpreted.

Protein sequencing reveals only 393 amino acid residues in the polypeptide chain of FDH while the full length gene accounts for 400 residues. This discrepancy cannot be resolved from the X-ray structure because the last residues in the models are Glu391 in hFDH and Ala374 in aFDH. The exact size of the mature FDH in solution therefore remains undefined as at most 391 residues are visible in the electron density, possibly due to a highly disordered conformation of the C terminus. The C-terminal seven residues would lie on the surface of the molecule and be highly accessible to solvent. In addition they contain three lysine residues. This would make them highly susceptible to proteolysis. Such post translational digestion of the C-terminal residues is thought to account for the multiple forms of FDH identified by isoelectric focusing. FDH obtained through cloning and expression in *Escherichia coli* is 400 amino acids long and coincides with one of the isoforms (Tishkov *et al.*, 1991).

The Ramachandran plot of the main-chain dihedral angles is presented in Figure 5. The most favoured regions shown in the Figure are based on an analysis of 118 structures (Morris *et al.*, 1992). A good quality model would be expected to have over 90% of the residues in these regions. The percentages are 89% for aFDH and 90% for hFDH. Ala198, participating in the FDH active centre, is in the disallowed zone and is discussed below.

Table 8

Differences in FDH sequences obtained from protein (Popov et al., *1990) and gene (Tishkov* et al., *1991)*

Number	Protein sequence	Gene sequence
77	Ser	Asp
137	Thr	Val
138	Val	Thr
145	Val	Cys
215	His	Val
216	Val	His
317	Asp	Asn

(b) *Hydrogen bond pattern within the subunit*

The FDH subunit is folded into two domains with the active centre cleft between them (Fig. 6). The domain structure has been described for the 3·0 Å hFDH analysis (Lamzin *et al.*, 1992). In brief, the co-enzyme binding domain is located in the middle of the FDH chain, (residues 147 to 333), and the catalytic domain is composed of the remaining residues. Both domains consist of a parallel β-sheet with left-handed twist surrounded by α-helices.

(i) *Secondary structure elements*

The FDH secondary structure is listed in Table 9 using nomenclature for the elements of the β-sheet in the co-enzyme-binding domain as historically suggested for dehydrogenases (Adams *et al.*, 1970). Residues are defined as helical if the carbonyl O from the first residue and the main-chain N-H of the last residue participate in a continuous H-bond pattern. For β-strands only those participating in a β-sheet are included. For β and π-turns the carbonyl O of the first residue forms an H-bond to the N of the last residue. H-bonds between main-chain carbonyl O and imide H within 1·3 to 2·5 Å were accepted. The distance between main-chain N and H atoms was assumed to be 0·86 Å. There are

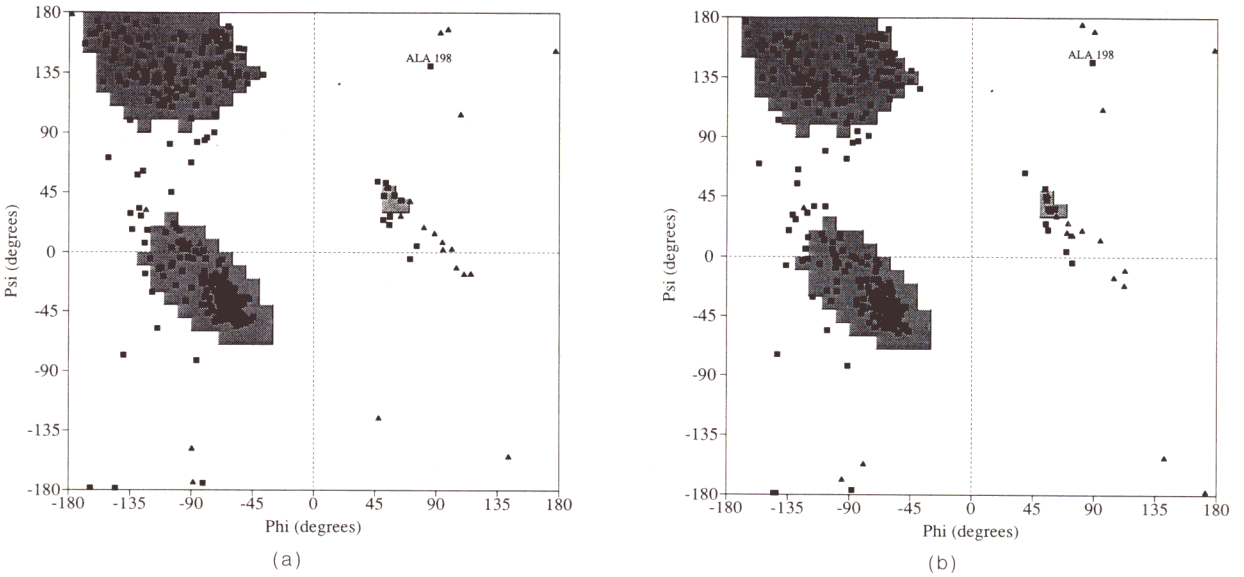


Figure 5. Ramachandran (ϕ/ψ) plot of (a) hFDH and (b) aFDH. Triangles correspond to glycine residues and squares to all other residues except proline. The most favoured regions (Morris *et al.*, 1992) are shaded. Ala198 is labelled as it is in the disallowed region.

substantial changes compared to the secondary structure based on hFDH at 3·0 Å resolution. Various 1–4, 1–5 and 1–6 contacts are now included.

α-Helices usually include both 4/13 (α) and 3/10 H-bonds. Several α-helices are completely transformed to 3/10 helices at their ends. Thus, in α1 the 4/13 H-bonds end at Ser65 N and there are 3/10 connections Leu63 O to Asn66 N and Glu64 O to Gly67 N. Similar patterns are observed in αB, α7, αE and αG. αD is a short helix with a 4/13 H-bond His258 O to Glu262 N and 3/10 bond His258 O to Lys261 N. There are three new pure 3/10 helices: left-handed helix 1A and right-handed helices FA and 9A. Both aFDH and hFDH models have the

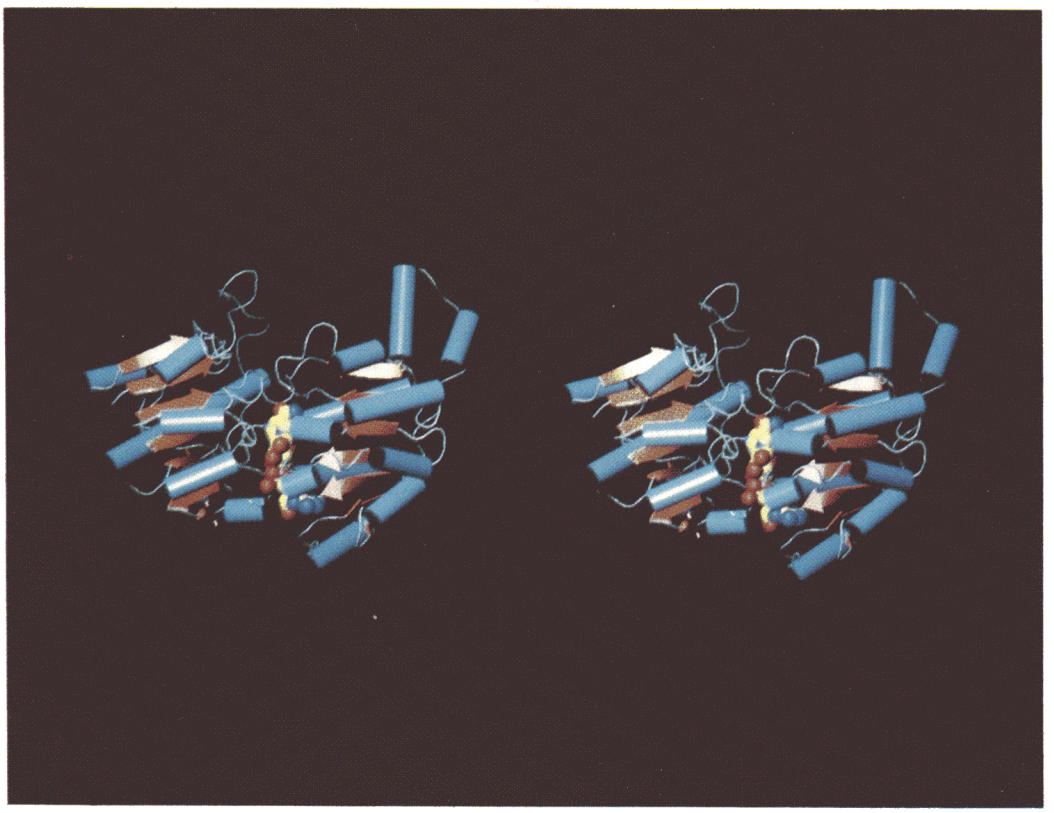


Figure 6. Stereo diagrammatic representation of the hFDH subunit with NAD bound.

Table 9
The secondary structure elements

Residue numbers	Element		Comments
1–7	β-Strand 1		
11–14	β-Turn I	*	
30–33	β-Turn II	*	Gly32 with positive ϕ
45–48	β-Turn II	*	Gly47 with positive ϕ
52–55	β-Turn I	*	
55–59	Left-handed 3/10 helix 1A	*	Glu56, Leu57, Gly58 with positive ϕ
59–67	α-helix 1		59–65 α-helix, 63–67 3/10 helix
63–68	π-Turn	*	Gly67 with positive ϕ
69–73	β-Strand 4		
81–87	α-Helix 2		
87–90	β-Turn I	*	
91–95	β-Strand 5		
96–99	β-Turn I	*	
104–111	α-Helix 3		
111–114	β-Turn I	*	
115–120	β-Strand 7		
124–127	β-Turn I	*	
128–136	α-Helix 4		
132–137	π-Turn	*	Asn136 with positive ϕ
138–140	β-Strand 8		
142–145	β-Turn II	*	Tyr144 with positive ϕ
146–163	α-Helix A		
160–165	π-Turn	*	Asn164 with positive ϕ
165–175	α-Helix 5		
171–176	π-Turn	*	Gly175 with positive ϕ
178–184	α-Helix 6		
183–186	β-Turn I	*	
189–192	β-Turn II	*	Ala191 with positive ϕ
192–198	β-Strand A		
200–214	α-Helix B		200–211 α-helix, 210–214 3/10 helix
216–220	β-Strand B		
226–234	α-Helix C		
230–235	π-Turn	*	Asn234 with positive ϕ
236–238	β-Strand C		
241–248	α-Helix 7		241–245 α-helix, 244–248 3/10 helix
249–254	β-Strand D		
258–262	α-Helix D	*	
261–264	β-Turn II	*	His263 with positive ϕ
266–273	α-Helix E		266–270 α-helix, 269–273 3/10 helix
274–277	β-Turn II	*	Gly276 with positive ϕ
278–281	β-Strand E		
284–288	3/10 Helix FA	*	
289–300	α-Helix F		
296–301	π-Turn	*	Gly300 with positive ϕ
304–307	β-Strand F		
316–319	β-Turn I	*	
319–324	α-Helix G		319–322 3/10 helix, 320–324 α-helix
324–327	β-Turn II	*	Tyr326 with positive ϕ
326–330	β-Strand G		
333–336	β-Turn I	*	
337–358	α-Helix 8		
354–359	π-Turn	*	Gly358 with positive ϕ
362–366	3/10 Helix 9A	*	
367–369	β-Strand 9	*	
372–374	β-Strand 10	*	
374–377	β-Turn II′	*	Gly375 with positive ϕ
375–381	α-Helix 9		

Those marked with * are new compared to 3·0 Å hFDH (Lamzin *et al.*, 1992). See the text for definition of particular elements and their boundaries.

same secondary structure elements except α9 located at the C terminus, which is disordered in the aFDH structure.

Turns are classified according to Wilmot & Thornton (1990). β-Turn type I is a fragment of a 3/10 helix with one H-bond only between O_i and N_{i+3}. Residues X_{i+1} and X_{i+2} have (ϕ/ψ) torsion angles typical for an α-helix and a 3/10 helix, respectively. In β-turn type II the X_{i+1} residue has torsion angles of the β region, while the X_{i+2} residue is in a left-handed helical conformation. β-turn type II′ is a "mirror image" of the type II β-turn and is consequently characterized by torsion angles of about (60°, −120°) for the X_{i+1} residue and an α-helical conformation for the X_{i+2} residue. According to Wilmot & Thornton (1990) the

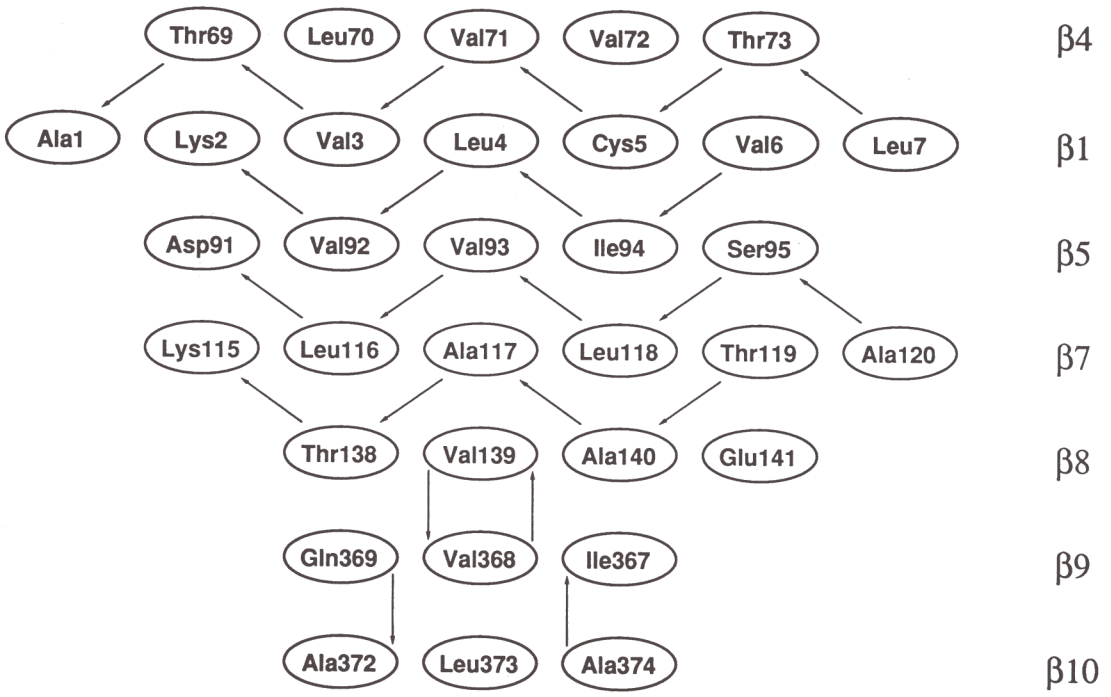


Figure 7. Topology of the β-sheet of the catalytic domain.-β-Strands are labelled according to Table 9. Pointers show H-bonds and their directions correspond to those from N to O.

frequency of occurrence of β-turns of types I, II and II′ are 47, 16 and 3%, respectively. In FDH there are nine, seven and one turn of the respective types and no other kinds of β-turn are found.

The 4·4/16 contact making an H-bond O_i to N_{i+5} is called a π-turn. This should not be confused with the π-helix that is fully composed of residues with torsion angles on average ($-57°$, $-70°$) close to the right-handed α region. The π-turn is usually characterized by dihedral angles close to the α-helix region for residues X_{i+1}, X_{i+2} and X_{i+3} and to the left-handed helical region for X_{i+4}. All seven π-turns are located at the ends of helices. All seven β-turns of type II have their X_{i+2} residues and all seven π-turns have their X_{i+4} residues in the left-handed helical conformation.

There are several changes in the definition of the β-strands compared to the 3·0 Å model. β-Strands previously defined as 2, 3, 6, 9, 10 and 11 are omitted as they belong neither to the β-sheet of the catalytic domain nor to that of the co-enzyme binding domain. They do not form any additional β-sheet with each other. Two new β-strands numbered β9 and β10 are included and extend the β-sheet of the catalytic domain (Fig. 7). β4, β1, β5, β7 and β8 comprise a parallel core, but the two extra strands make an anti-parallel part of the sheet. Thus, the β-sheet of the catalytic domain is composed of an extended parallel part and a small antiparallel region.

(ii) *Non-glycine residues with positive ϕ angles*

Glu56 and Leu57 belong to a left-handed 3/10 helix 1A. Two H-bonds stabilize this helix. In addition Leu57 N makes an H-bond to Leu50 O and the Glu56 side-chain is stabilized by a salt bridge with the Arg60 side-chain. Residues Asn136, Asn164 and Asn234 are X_{i+2} residues in type II β-turns while Tyr144, Ala191, His263 and Tyr326 are X_{i+4} residues in π-turns. The left-handed helical conformation of each of these residues provides an additional H-bond making it energetically favourable. The reason for the left-handed helical conformation for residues Asn178 and Asp214 is not so obvious. If the peptide bond between residues 177 and 178 is flipped to provide Asn178 with a negative ϕ angle, this would cause a reorientation of the Trp177 side-chain making it overlap with Pro331. The latter seems to be important in stabilizing the extra βG strand in the co-enzyme binding domain compared to other dehydrogenases. A different explanation can be given for Asp214. It is the last residue in helix B and participates in the H-bond pattern of the helix, so the spatial positions of its N and CA atoms are fixed. The next strand βB starts from residue 216 so the chain fold from 214 to 216 must be fixed. Asp214 must therefore be in the left-handed helical conformation.

In both aFDH and hFDH Ala198 has the same "disallowed" conformation. All NAD/FAD-dependent enzymes for which structures are available to date contain glycine in this position. It was noted (Lamzin *et al.*, 1992) that the presence of Ala198 in FDH disproved previous assumptions about the essential need for glycine at this position (Rossmann *et al.*, 1975) and the fingerprint pattern with the conserved motif Gly-X-Gly-X-X-Gly, where X is any amino acid (Wierenga *et al.*, 1985; Scrutton *et al.*, 1990). The high resolution analysis confirms the presence of alanine at position 198. The electron density both for the aFDH and hFDH models in both subunits is typical for an alanine

residue and contains clear density for its CB atom. The abnormal conformation of Ala198 (normal for glycine) is stabilized by three additional H-bonds. Ala198 belongs to the β-sheet of the co-enzyme binding domain at the end of the βA strand. Firstly within the β-sheet Ala198 N makes an H-bond with Thr220 O from βB. Secondly Ala198 O makes an H-bond with Thr196 OG1. Finally Ala199 N, whose location is fixed by the carbonyl group of Ala198, makes an H-bond to Asp221 OD2. Proper orientation of the Asp221 side-chain is essential for binding NAD (Wierenga *et al.*, 1985; Lamzin *et al.*, 1992). The main-chain dihedral angles for Ala198 correspond to those found for the first glycine in the fingerprint (Wierenga *et al.*, 1985). Thus, such a conformation is important for maintaining the correct three-dimensional conformation of the Gly/Ala-X-Gly-X-X-Gly motif. The first residue does not need to be a glycine but it is certainly the preferred residue at this position. In formate dehydrogenases from yeast (Hollenberg & Janowicz, 1989) and potato (Colas des Francs, personal communication) this position is occupied by glycine.

(c) *Molecular symmetry*

The two subunits of the FDH dimer were superimposed using all atoms and are related by an almost perfect non-crystallographic 2-fold symmetry axis. In hFDH the rotation angle is 179·9° and the r.m.s. deviation in spatial position between the subunits is 0·19 Å for CA atoms, 0·11 Å for NAD atoms, 0·22 Å for azide atoms and 0·47 Å for all atoms. In aFDH the rotation angle is 179·2° and the deviations for CA and all atoms are 0·36 and 0·59 Å respectively.

r.m.s. differences in CA positions after superposition are plotted as a function of residue number in Figure 8. The aFDH model is less symmetrical than hFDH. The highest deviations from symmetry generally correspond to β or π-turns. However, in several places they occur in other elements. In hFDH the αC helices in the co-enzyme binding domain (residues 226 to 233) deviate about 1 Å from the symmetry. In spite of the fact that these helices are present in all other dehydrogenases, contribute to the intersubunit interactions and are approximately structurally conserved, in FDH they are located at the periphery and do not participate in making contacts with other structural elements. Extra flexibility for these helices may arise from atypical intersubunit interactions in FDH making it a "P-dimer" (Lamzin *et al.*, 1992). However, in aFDH these helices follow the overall symmetry well. The aFDH model shows a deviation from ideal 2-fold symmetry for a number of other α-helices: the short αD-helix in the co-enzyme binding domain, located at the molecular surface in aFDH and α2, α3 and α4-helices from the catalytic domain. The latter surround the β-sheet of the catalytic domain and are accessible to solvent, a possible reason for their asymmetry. The opposite part of the β-sheet "sandwich" consists of helices α1, α8 and 9A, additionally stabilized by a long loop Tyr30 to Gln48 covering these helices and following the molecular symmetry closely.

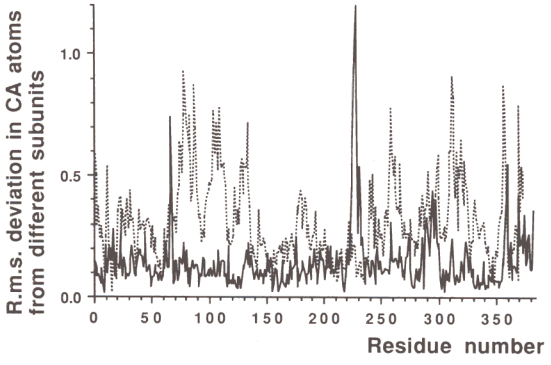


Figure 8. r.m.s. deviation in CA atom positions when 2 subunits are superimposed. Continuous line, hFDH model; broken line, aFDH model.

In spite of the very limited intermolecular contacts, which in neither structure extend to the regions where the symmetry is violated, deviations from ideal 2-fold symmetry within the dimer must be related to crystal packing. Lack of substantial deviation from symmetry in the dimer is in keeping with the known independence of its two active centres (Popov *et al.*, 1978) and absence of the "half-of-the-sites reactivity" phenomenon observed for a number of dehydrogenases (Hodges *et al.*, 1978).

Many water molecules follow the molecular symmetry. A total of 354 out of 788 solvent molecules in hFDH are related by the molecular 2-fold axis, with less than 0·5 Å deviation and 238 waters less than 0·3 Å. In aFDH there are 213 and 110 out of 854 such water molecules. The most important symmetrically related water molecules are those located near the enzyme active centre, as discussed below.

8. Comparison of Apo and Holo FDH

(a) *NAD binding*

NAD binds to the FDH active centre in the elongated conformation typical for other dehydrogenases (Eklund & Bränden, 1987). The distance C2N to C6A is 14·7 Å and the angle N9A-OPP-N1N is 108° in both subunits. NAD can be represented as a set of mutually perpendicular planes. Adenine and ribose moieties are essentially perpendicular to each other. The nicotinamide plane is approximately parallel to the adenine ribose.

The location of NAD in the active centre of hFDH is shown schematically in Figure 9. It makes a number of H-bonds with protein and these are considered separately for each part of NAD. Water molecules included in the Figure occupy the same positions in both subunits and the r.m.s. deviation from symmetry in their positions is 0·3 Å.

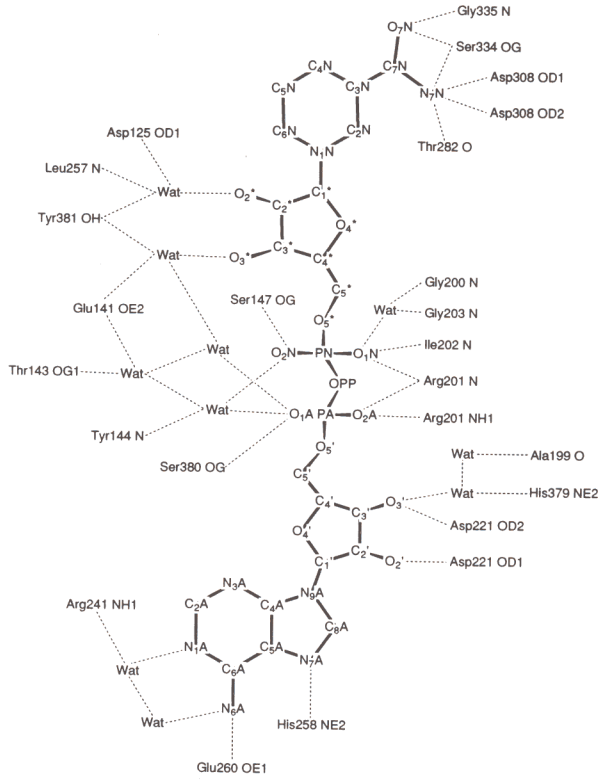


Figure 9. Schematic representation of NAD bound to the active centre. H-bonds from NAD to protein atoms either direct or through water molecules are shown by broken lines.

Table 10

Distances (Å) between the carboxamide moiety of NAD and neighbouring residues in hFDH compared to the hypothetical model where nicotinamide is perfectly planar.

		Hypothetical model with planar nicotinamide	hFDH model
A. *Distances to atoms participating in hydrogen bonding with NAD*			
NAD N7N	Thr282 O	2·8	2·8
NAD N7N	Asp308 OD1	3·4	3·1
NAD N7N	Asp308 OD2	3·5	3·2
NAD N7N	Ser334 OG	3·5	3·0
NAD O7N	Ser334 OG	2·7	3·1
NAD O7N	Gly335 N	3·1	3·2
B. *Distances to atoms not participating in hydrogen bonding with NAD*			
NAD N7N	Val309 CG1	3·4	3·7
NAD N7N	His332 NE2	2·9	3·2
NAD O7N	Val150 CG2	3·5	3·6
NAD O7N	His332 NE2	3·1	3·1

Discrepancy in distance of more than 0·3 Å is significant (see the text).

(i) *Nicotinamide*

The carboxamide group of this unit is connected by six H-bonds with main and side-chain atoms of four residues: O7N to Ser334 OG (3·1 Å) and to Gly335 N (3·2 Å), N7N to Thr282 O (2·8 Å), Asp308 OD1 (3·1 Å), OD2 (3·2 Å) and to Ser334 OG (3·0 Å). In aFDH the positions corresponding to O7N and N7N are occupied by two water molecules. The residues participating in the binding are from the co-enzyme binding domain (ends of βE and βF) and from the β-turn between βG and α8 connecting the two domains. The carboxamide group is twisted by 25° from the pyridine plane with O7N facing C4N. A comparison of distances between the carboxamide group and neighbouring atoms is shown in Table 10. Two models are considered: the nicotinamide unit as in hFDH and a hypothetical model with a perfectly planar nicotinamide superimposed on that in the holo structure. The error in co-ordinates in hFDH is estimated from the σ_A plot to be 0·11 Å giving the error in interatomic distances of 0·16 Å. Hence, the difference between distances measured twice (in the A and the B subunits) exceeds 1σ if it is more than 0·16 Å. Thus, discrepancies in distances listed in Table 10 between the hypothetical and real models of more than 0·3 Å (2σ) are statistically significant. The NAD N7N and O7N atoms acquire H-bonds with Asp308 and Ser334 side-chains in the hFDH model compared to the hypothetical model with in-plane nicotinamide. The non-bonding distance N7N to Val309 CG1 becomes shorter in the "planar" model. The His332 side-chain points towards the carboxamide C7N, its imidazole plane is perpendicular to the carboxamide plane and this residue is also considered to have a non-bonded interaction. In hFDH the shortest distances between His332 NE2 and carboxamide atoms are about 3·2 Å. In the "planar" model the distance His332 NE2 to N7N is shortened to 2·9 Å. Thus, the carboxamide binding site requires the carboxamide plane to be non-planar with the pyridine.

(ii) *Nicotinamide ribose*

All interactions with nicotinamide ribose are made through water molecules. Atom O2* makes an H-bond to water and then to Leu257 N (loop connecting βD and αD in co-enzyme binding domain) and to side-chains Tyr181 (α6 in the co-enzyme binding domain) and Asp125 (β-turn between β7 and α4 in the catalytic domain). An equivalent water molecule is present in aFDH but it is shifted by about 1 Å from the position of NAD O2*. O3* forms H-bonds through waters to Tyr181 and Glu141 (loop connecting β8 and αA at the interdomain contact). An H-bond to Tyr181 is formed in hFDH while in aFDH this residue is disordered. This, and other residues from the C terminus which interact with NAD in hFDH, are discussed below in section (b).

(iii) *Pyrophosphate*

The pyrophosphate part of NAD lies close to residues Ala198 to Gly203 belonging to the Gly/Ala-X-Gly-X-X-Gly fingerprint. The αB helix starts from Gly200, at the third position in this pattern. The main-chain nitrogen atoms of the first four residues (Gly-X-X-Gly) are believed to provide

a positive charge to bind the pyrophosphate moiety in a number of dinucleotide binding enzymes. There is an H-bond between NAD O1N and N of the third X residue (Wierenga *et al.*, 1985). In FDH O1N forms two H-bonds with main-chain atoms of αB: Arg201 N (3·3 Å) and Ile202 N (3·0 Å), residues 2 and 3 in the helix. Interactions between O1N and other residues from αB are formed through a water molecule to Gly200 N and to Gly203 N. There is a water molecule in aFDH occupying the same position. O2A is involved in binding to αB by making an H-bond with Arg201 N (3·1 Å). Most of structurally known dinucleotide binding enzymes have an arginine residue to further neutralize the negative charge on the pyrophosphate (e.g. see Wierenga *et al.*, 1985). Only in FDH is such an arginine located at the beginning of the same αB helix. Its side-chain is bent to make an extra H-bond to O2A (2·8 Å). The other side of the pyrophosphate unit points towards the loop connecting the catalytic and co-enzyme binding domains. O2N makes an H-bond to Ser147 OG (2·8 Å) and O1A binds to Ser380 OG (2·7 Å), while the latter is disordered in aFDH. O2N and O1A also interact with two water molecules and through those to the Glu141 and Thr143 side-chains and Tyr144 N. A similar H-bond pattern supported by three water molecules is found in aFDH. Thus, the environment of pyrophosphate is highly hydrophilic, positively charged and, except Ser380, comprises residues from the co-enzyme binding domain and the interdomain region.

(iv) *Adenine ribose*

Two ribose hydroxy oxygen atoms, O2′ and O3′, are connected by H-bonds to Asp221 OD1 and OD2 both with distances of 2·7 Å. Such an aspartate residue is always present in NAD-dependent enzymes, while FAD binding proteins have a glutamate at this position (Wierenga *et al.*, 1985). This residue is responsible for recognition of NAD and its spatial position (end of βB) is highly conserved. There is no such residue in NADP binding enzymes. O3′ also interacts through water molecules with Ala199 O, located between βA and αB, and with His379 NE2. The H-bond between water and His379 appears only in hFDH. In aFDH the Asp221 side-chain makes H-bonds to two water molecules, occupying approximately the positions of O2′ and O3′.

(v) *Adenine*

All residues around this part of NAD belong to the co-enzyme binding domain. Adenine lies in the extensive hydrophobic pocket made up by the side-chains of Val197, Leu253, Cys255, Thr261, Met264 and Leu287 (not shown in Fig. 9). This hydrophobic core is formed by residues from the β-sheet of the co-enzyme binding domain and from the αD, αE and FA helices shielding the sheet from the solvent. An area on the opposite side of the adenine is hydrophilic and a few H-bonds are found: N7A to His258 NE2 (3·1 Å), N6A to Glu260 OE1 (3·0 Å), N6A and N1A through water molecules to Arg241 NH1.

Cys255 has been identified as an "essential thiol", critical for maintaining activity and stability on the basis of selective modifications with radioactive and fluorescent labels (Ustinnikova *et al.*, 1988). The high resolution structures provide an explanation of the behaviour of this cysteine. In hFDH it is located just above the adenine plane of the NAD. In aFDH this cysteine is easily accessible to the modifying agents. Introduction of any bulky substituent in place of the SH group as a result of modification or simple oxidation catalysed by transition metal ions would result in steric hindrance between this substituent and adenine and thus prevent productive binding of NAD. In hFDH Cys255 is protected by the bound co-factor and inaccessible to the modifying agents.

The majority of direct (not through water molecules) H-bonds between NAD and protein are formed by residues from the co-enzyme binding domain. Conformations of residues responsible for NAD binding show only minor changes between aFDH and hFDH. Several water molecules in the aFDH active centre, occupying the space where NAD is to be bound, are responsible for maintaining important residues from the co-enzyme binding domain in the active conformation. Thus, the residues from the co-enzyme binding domain are approximately fixed in aFDH so as to be correctly oriented for binding NAD in the proper conformation and place.

(b) *Domain movement*

The structures of hFDH and aFDH differ mostly in the location of the peripheral catalytic domains (Fig. 10). A movement of these domains from the position in apo to that in holo was proposed to be one of the most important features of enzyme catalysis by NAD-dependent dehydrogenases (Eklund & Bränden, 1987). FDH fits into this general scheme. The hFDH and aFDH models were compared by superimposing the internal co-enzyme binding domains, which pack against one another around the 2-fold axis of the dimer. The overall r.m.s. deviation in CA atoms for these internal domains is 0·49 Å. Deviations are substantially higher for residues before 146 and after 341, i.e. in the region of interdomain contacts or in the catalytic domain.

The movement of the catalytic domains in hFDH compared to aFDH is essentially a rotation by about 7·5° around hinges connecting residues 146–147 and 340–341. Surprisingly these hinges are located not within irregular loops or turns but in the αA and α8 helices, respectively. No substantial differences in main-chain dihedral angles between aFDH and hFDH models are found near these hinges. Helices αA and α8 are the longest in the structure (Table 9) and both begin at the interdomain contact. αA penetrates deeply into the co-enzyme binding domain, interacts with αA from the other subunit and makes up the core of the FDH dimer. α8 runs through the catalytic domain, shields the β-sheet of the catalytic domain and ends at the

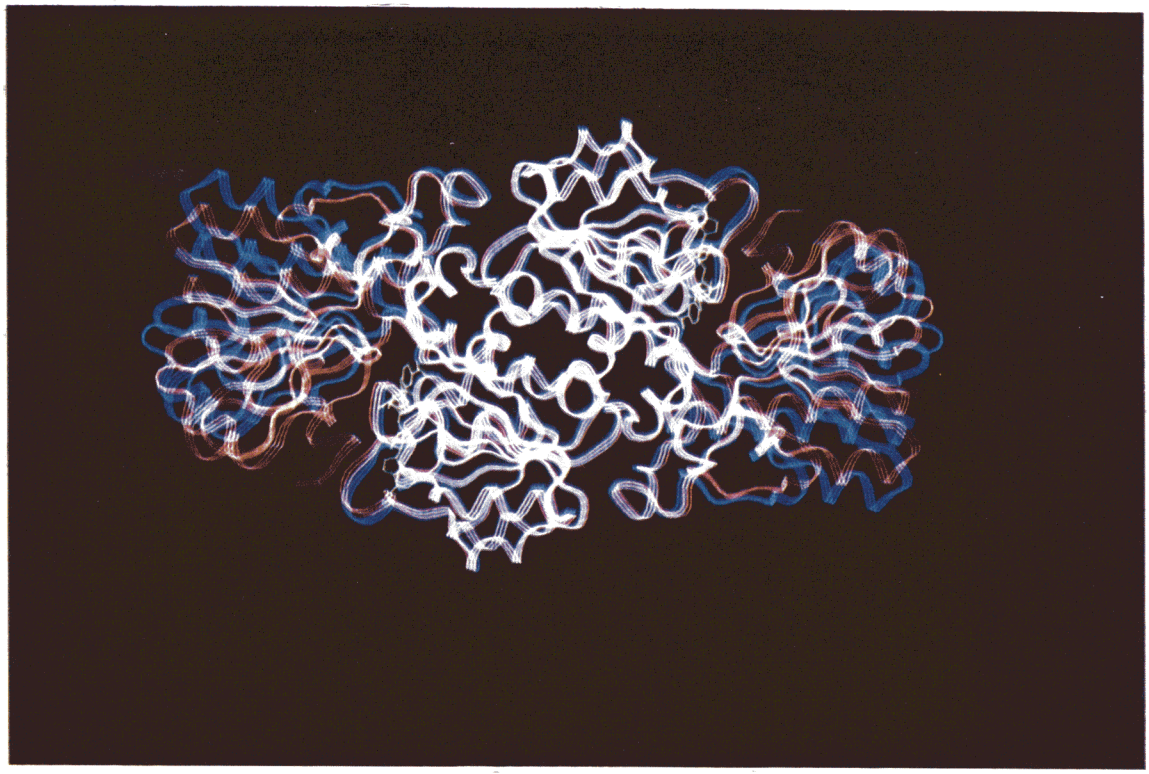


Figure 10. Ribbon plot of hFDH (red) and aFDH (blue) models. NAD is shown in green. The superposition was made using residues 147 to 333, making up the internal co-enzyme binding domains, in both subunits. The 2 models differ most in the location of the peripheral catalytic domains.

far side of the protein. These helices seem to be important in making the interdomain contacts relatively rigid and allowing only small well-defined rotations to occur. The rotation of the catalytic domain results in only a minor distortion of these helices at their ends.

The rotation of the catalytic domain occurs so that in hFDH this domain has moved essentially as a rigid body nearer to the active centre shielding it from solvent (Fig. 10). Similar changes take place in other NAD-dependent dehydrogenases, where the apo form was termed as having an open shape, while in the holo the protein acquires a closed conformation (e.g. see Grau *et al.*, 1980). Different kinds of domain rotation have been described (Ptitsyn *et al.*, 1988). For a "sliding" movement the protein shape remains approximately intact, while in a "rolling" movement, as in FDH, the overall shape changes and the conformational alterations can sometimes be detected by less detailed methods of structure determination, such as small angle X-ray scattering (Lamzin *et al.*, 1986).

Figure 11(a) shows NAD from the hFDH model superimposed on aFDH. NAD lies on the surface of the co-enzyme binding domain and there is a large cleft between the co-enzyme and the catalytic domain. Figure 11(b) shows NAD in hFDH where the active centre is completely closed by the catalytic domain and is not accessible. Therefore, NAD seems to bind to aFDH in the open conformation followed by domain movement and closing of the active centre. A scheme, where NAD binds to an active centre, already in the closed conformation, is unlikely to occur because of steric hindrance.

Thus, we can assume that the rotation of the catalytic domain occurs after the NAD molecule has bound to the active centre. The main question is what contacts NAD makes with the protein to cause domain movement? Important interactions between NAD and protein are to be found, which initiate further conformational changes and bring the catalytic domain to the position where the reaction occurs. There are several observations suggesting that these interactions originate from the co-enzyme binding domain.

After NAD is bound the residues from the co-enzyme binding domain responsible for this binding only slightly adjust their conformation. However, this results in subsequent changes in the co-enzyme binding domain and even more in the catalytic one. The side-chain of Asp221 rotates around its CB–CG bond by about 60° to optimize two H-bonds with atoms O2′ and O3′ of the adenine ribose. Asp221 OD1 is now able to make an H-bond with His223 N, which moves slightly towards the Asp221 side-chain. As a result the distance Asp221 OD1 to His223 N becomes 2·7 Å in hFDH compared to 4·5 Å in aFDH. Formation of this H-bond and shift of the main-chain atoms of His223 causes movement of its side-chain by about 3·0 Å. The interactions of the individual parts of NAD bound to the co-enzyme binding domain with residues from the

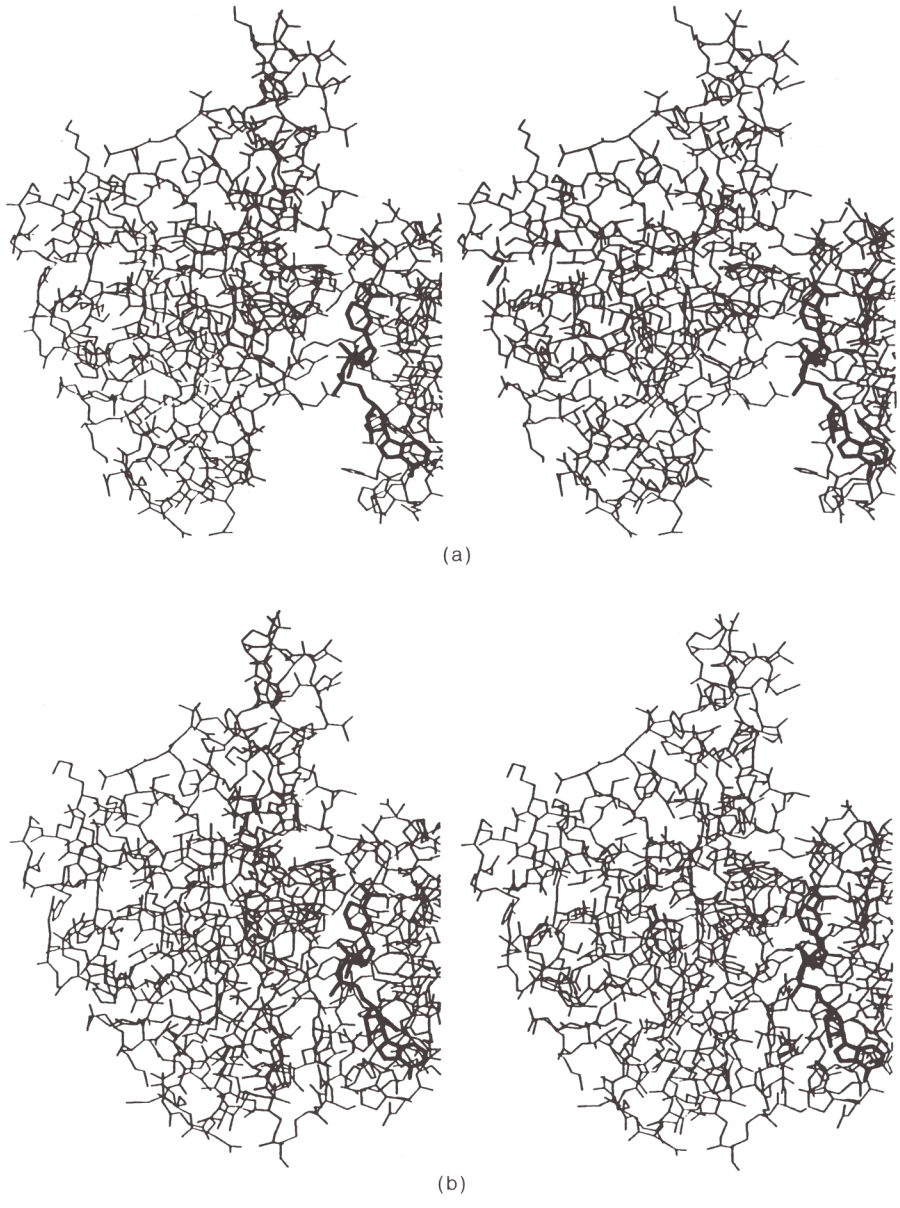


Figure 11. Stereo view of a region of the FDH subunit showing the movement of the catalytic domain, closing the active centre. The catalytic domain is on the left and a part of the co-enzyme binding domain is on the right. (a) First stage of acquiring the catalytic conformation: aFDH with NAD, placed as it is bound in hFDH. (b) Second stage: hFDH with NAD. The co-enzyme binding domain and NAD are in the same positions in (a) and (b).

catalytic domain cause substantial changes in the latter which are now discussed.

(i) *Adenine ribose*

The movement of His223 caused by interactions with the adenine ribose results in a new H-bond between His223 NE2 and His379 O. O3′ of the adenine ribose forms an H-bond through water to His379 NE2. Thus, the conformation of His379 in hFDH differs from that in aFDH and is stabilized by two new H-bonds.

(ii) *Adenine*

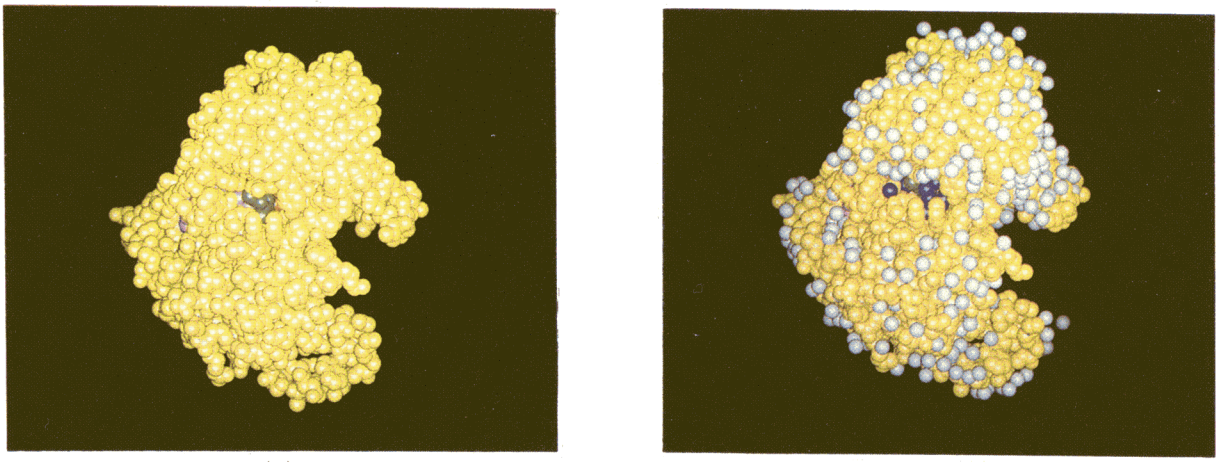
Residues Glu260 and Arg222 from the co-enzyme binding domain change their conformation and make an H-bond Glu260 OE1 to Arg222 NH2 (3·2 Å). Arg222 NH1 is now able to form new H-bonds with Tyr381 O and His379 O from the catalytic domain. In hFDH His258 becomes ordered and its NE2 makes an H-bond to adenine N7A and His258 ND1 forms an H-bond to Ser382 OG.

(iii) *Pyrophosphate*

O1A of the pyrophosphate group forms a direct H-bond with Ser380 OG.

(iv) *Nicotinamide ribose*

Atom O3* of the nicotinamide ribose makes an H-bond to a water molecule and then through it to Tyr381 OH.



(a) (b)

Figure 12. The substrate binding channel in the hFDH subunit. The van der Waals representation of residues Ala1 to Lys383 is shown in yellow; the 2 side-chains, Asp125 and Arg284, shielding the channel in green; NAD in magenta; azide in red; 10 water molecules located in the channel in blue; other water molecules in cyan. (a) Only protein, NAD and azide atoms are shown. The atoms corresponding to the NAD and azide in the active side are accessible to solvent provided the Asp125 and Arp284 side-chains are flexible. (b) Solvent molecules are included.

(v) *Nicotinamide*

This part of NAD forms direct or indirect H-bonds with residues from the catalytic domain which do not change their conformation in hFDH compared to aFDH.

Thus, after NAD is bound to the active centre and properly oriented by residues from the co-enzyme binding domain (Fig. 11(a)), new interactions with the catalytic domain are formed. Residues His379, Ser380, Tyr381 and Ser382 acquire additional H-bonds. All these residues belong to the C-terminal α9 helix. The region after Ala374 is completely disordered in aFDH. It becomes ordered only after NAD is bound and provides additional H-bonds to stabilize this region directly or through water molecules. This stabilization produces further changes. Helix 9A, preceding α9, is shifted by about 4 Å in hFDH compared to aFDH. Helix 9A is located about 20 Å away from the hinges of the domain rotation and such a 4 Å movement is easily achieved, analogous to the use of a lever for moving heavy objects by the well-known mechanical principle "give me but one firm spot on which to stand, and I will move the earth" (Archimedes, 230 BC). The formation and ordering of the C-terminal α9 in hFDH results in the movement of all neighbouring fragments and the overall rotation of the catalytic domain (Fig. 11(b)). There are other minor additional interactions, either hydrophilic or hydrophobic, between NAD and the catalytic domain, which may be important. Nevertheless the main force of the domain movement appears to be that part of the NAD binding energy which is transformed into a constraint of the protein structure through the ordering of the C terminus. The minor adjustment of the Asp221 side-chain may assist this avalanche of interactions resulting in the enzyme taking up the catalytic conformation. A possible role of azide (or substrate) in accomplishing conformational changes in FDH awaits further elucidation.

(c) *Substrate channel*

The active site lies well below (about 15 Å) the surface of the protein. In both hFDH and aFDH there is a wide channel running from the active site to the surface. Two side-chains, Asp125 and Arg284, lie close to the active site in the bottom of the channel and divide it in two parts. Azide bound to the active centre, viewed from the solvent is seen on the right (Fig. 12(a)). In hFDH the channel is additionally shielded from solvent by the loop of residues Asn385 to Ser390, which is ordered in one subunit only. Three side-chains, Pro97, Arg284 and His332, form the right half of the bottom of the channel and delineate the gateway for the substrate to proceed to the active centre. This bottleneck allows access to the active site from the solvent provided the Arg284 side-chain is flexible.

In hFDH the channel is filled by ten water molecules, (Fig. 12(b)). The positions of these water molecules are the same in both subunits deviating by less than 0·2 Å from the molecular 2-fold symmetry. All ten water molecules are connected by H-bonds, characterized by O to O distances of 2·7 to 3·0 Å and interbond angles 93° to 134°, typical for tetrahedral co-ordination. In aFDH there are six water molecules in the channel occupying the same positions as in hFDH. Thus, the channel may provide the pathway for substrate when NAD is already bound to the active centre. The channel seems to be large enough to accommodate a molecule bigger than formate but not NAD. There are two routes through which formate may reach the active site: through the NAD binding area if NAD is absent or through the channel if NAD is bound.

(d) *The active site*

It is now possible to discuss the structure of the active centre around the catalytic point (C4N of the nicotinamide unit of NAD). Only conclusions from the crystal structures are presented here. A discussion of the enzyme properties in solution, i.e. inhibition studies, pH-dependencies of kinetic parameters, identification of essential amino acid residues, coupled with knowledge of its tertiary structure and site-directed mutagenesis experiments, requires additional study and will be published elsewhere.

It is necessary to estimate the location of the formate anion in the active centre. hFDH is the ternary complex enzyme–NAD–azide. Azide is a potent inhibitor of FDH (K_i about 0·1 μM) competitive to formate (Popov *et al.*, 1978; Hermes *et al.*, 1984). The position of azide in the hFDH dead-end ternary complex can be taken as a first approximation to the formate binding site. When the structure of the transition state is considered, some minor changes in protein side-chain and nicotinamide conformations can be expected. Positions of azide in the dead-end and formate in productive ternary complexes may not be identical. Furthermore, as the structure of NADH bound to FDH is not known it is difficult to predict the exact position of formate, i.e. the distance NAD/NADH C4N to formate H and orientation of the formate C-H bond with respect to the nicotinamide. There are certain restrictions on the geometry of the transition state when hypothesizing substrate location and formulating the molecular mechanism of FDH. To allow efficient hydride anion transfer the distance between the formate H atom and C4N should be about 1·0 to 1·5 Å, the formate C-H bond should lie in the plane perpendicular to the NAD pyridine plane and passing through the C4N atom and the formate C-H bond should make angles of about 110° with the bonds C4N–C3N and C4N–C5N to allow the change of C4N from a planar to a tetrahedral configuration.

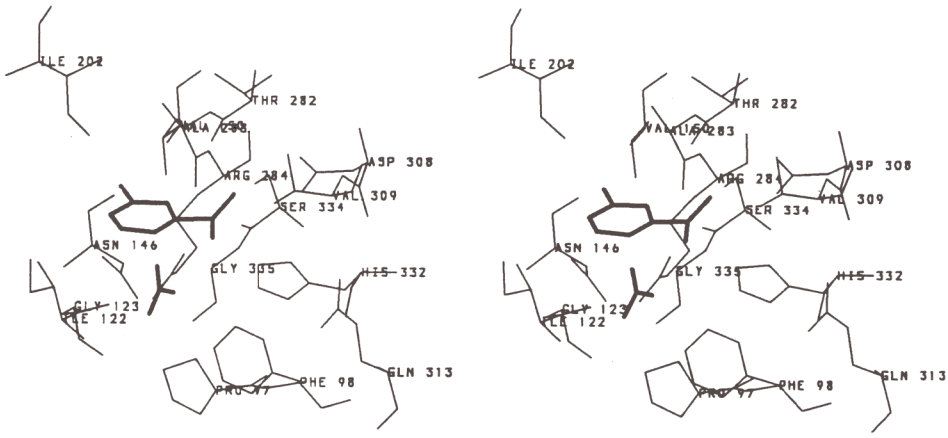
Taking into account the above constraints two alternative modes of formate binding can be suggested. In the first the carbon and two oxygen atoms of the formate are superimposed on the three atoms of azide and the C-H bond is oriented towards C4N. This gives a distance between C4N and formate H atoms of about 1·4 Å and angles for bonds formate C-H, C4N–C3N of 90° and formate C-H, C4N–C5N of 100°. The formate O1 (corresponding to azide N1) points towards Arg284 making an H-bond to its NH1 (3·2 Å) and NH2 (3·3 Å) atoms. The formate O2 (corresponding to the azide N3 atom) makes H-bonds to Asn146 ND2 (3·1 Å) and Ile122 N (2·9 Å). Viewed from the formate H the two oxygen atoms point to the Pro97 side-chain and the closest distance of 3·2 Å is the formate O2 to Pro97 CG. However, the negatively charged formate O1 is 2·8 Å from Ile122 O.

A second location of formate is obtained by moving it from the first to make angles with bonds C4N–C3N and C4N–C5N equal to 110° and rotating it to an H-bond distance between formate O1 and His332 NE2. However, the formate–histidine contact does not have the appropriate geometry for an H-bond as the distances formate O1 to both His332 NE2 and its hydrogen are 2·8 Å. The His332 side-chain must be rotated to make an H-bond to formate O1. In this location, formate only partially overlaps the azide binding site but is placed symmetrically around the plane perpendicular to the pyridine and passing through C4N. The formate O2 atom is at the same place as in the first location and makes the same H-bonds with Asn146 ND2 and Ile122 N. The formate oxygen atoms point between the side-chains of the hydrophobic Pro97 and Phe98 making multiple van der Waals contacts. An undesirable contact (2·7 Å) between formate O2 and Phe98 CE1 can be avoided by a movement of the Phe98 side-chain. As in the first location there is a short contact between two oxygen atoms: formate O1 to NAD O7N (2·8 Å).

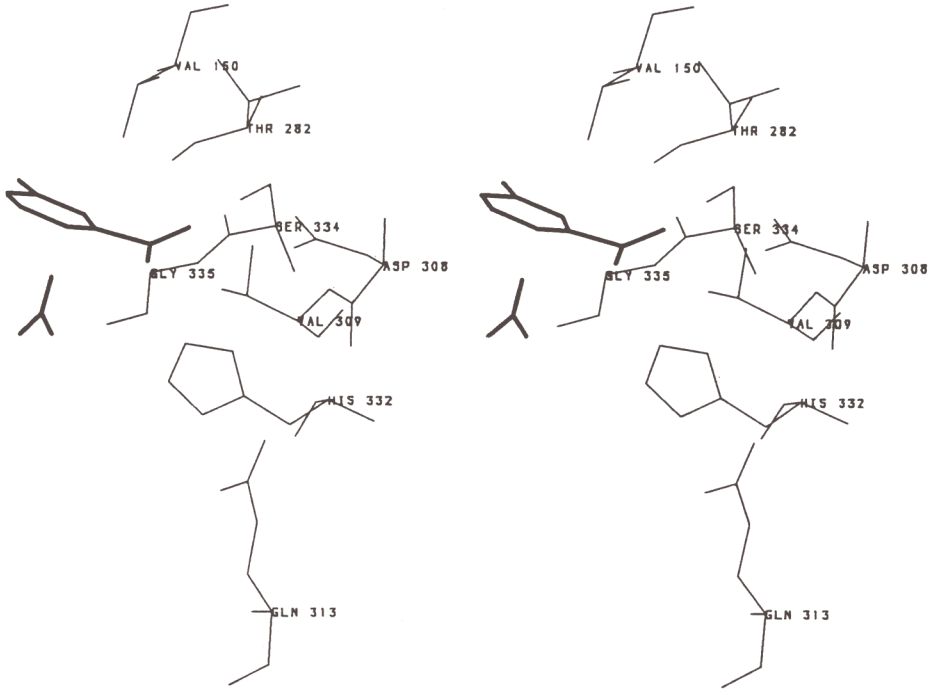
The difference between the two alternative locations proposed for the formate can be summarized as follows. In the first the formate is located between Arg284 and Asn146 and makes H-bonds with them. His332 does not participate in substrate binding. In the second the formate is located between His332 and Asn146 and the positive charge of the Arg284 side-chain is not compensated. To introduce an H-bond and to avoid a close contact the His332 and Phe98 side-chains must move. The resulting geometry of the complex is not ideal in either location. In both, catalysis occurs on the so called A-side (Rossmann *et al.*, 1975) of the nicotinamide. Relying on the structural data we believe the first location of formate is most likely to be its true position.

A stereo view and schematic representation of the active site are shown in Figures 13 and 14 with formate in the first location. The active site of FDH comprises negatively and positively charged clusters as well as hydrophobic patches. The "negatively" charged region, (Fig. 13(b)), is made up of residues Thr282, Asp308 and Ser334. These residues form several H-bonds with the carboxamide group of NAD and with each other. The role of this cluster is to hold the nicotinamide fragment in the proper conformation. Thr282, Asp308 and Ser334 occupy approximately the same positions and side-chain orientations in both aFDH and hFDH. They belong to the co-enzyme binding domain, are already properly oriented and "are waiting" for NAD to bind.

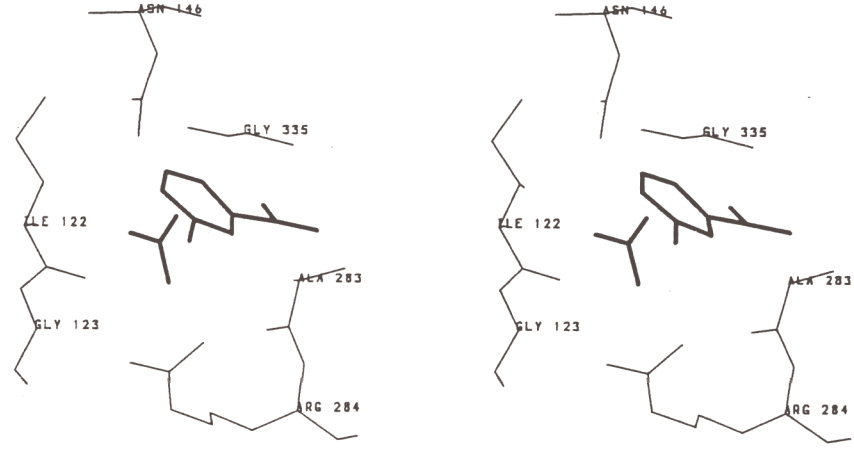
Another cluster of residues can be labelled as "positively" charged (Fig. 13(c)), and comprises Arg284, Asn146 side-chains and amide nitrogen of Ile122. In both locations of formate the Asn146 side-chain interacts with its O2 atom. The Asn146 side-chain orientation is supported by H-bonds ND2 to Gly335 O and OD1 to Ser147 N. The position of Asn146 in hFDH differs from aFDH by about 0·7 Å (Fig. 15), and thus acquires the appropriate conformation only after movement of the catalytic domain. Asn146 may also participate in


ILE 202
THR 282
ALA 283
ARG 284
ASP 308
VAL 309
SER 334
ASN 146
GLY 335
HIS 332
GLY 123
ILE 122
PRO 97
PHE 98
GLN 313

(a)

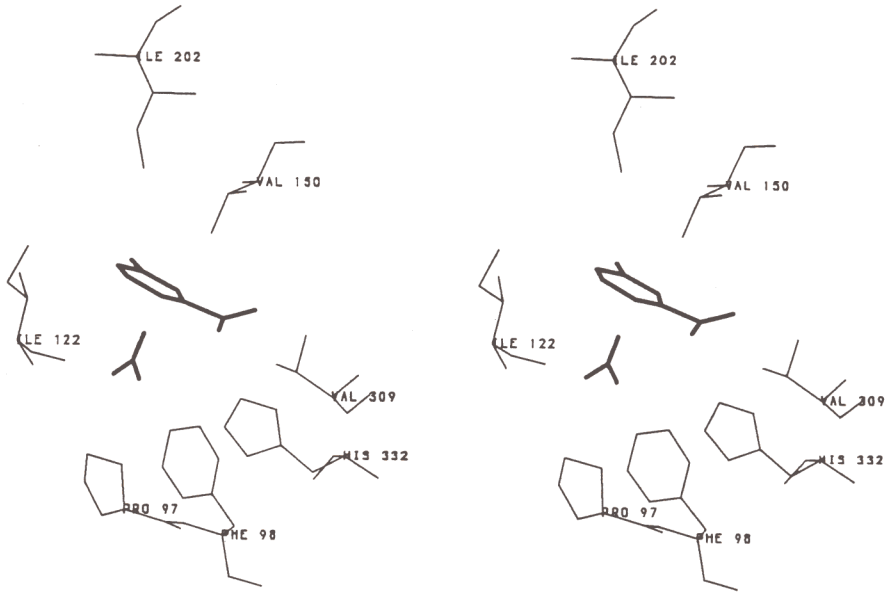

VAL 150
THR 282
SER 334
GLY 335
ASP 308
VAL 309
HIS 332
GLN 313

(b)


ASN 146
GLY 335
ILE 122
ALA 283
GLY 123
ARG 284

(c)

Fig. 13.



(d)

Figure 13. Stereo view of the FDH active centre near the point of catalysis (C4N atom of the nicotinamide unit of NAD). The location of the formate anion was derived from the position of the azide inhibitor and stereochemical restraints (see the text). NAD nicotinamide unit and formate (HCO_2) are shown in bold. (a) All residues surrounding the active site. (b) Carboxamide binding site with "negatively" charged cluster and several hydrophobic resides. (c) Formate binding site with "positively" charged cluster. (d) Hydrophobic areas around the active site.

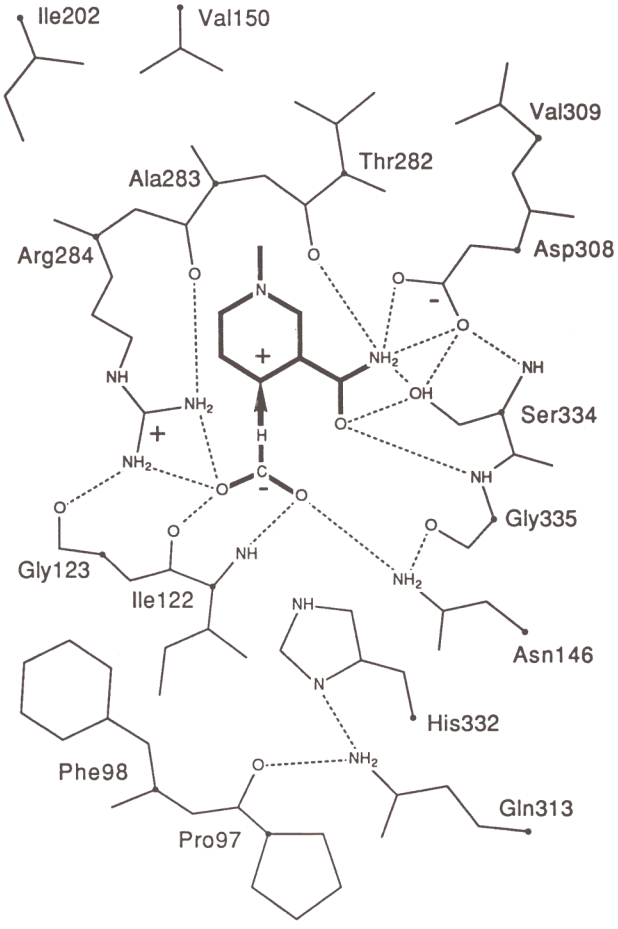


Figure 14. Schematic representation of the hydrogen bond network in the FDH active centre near the point of catalysis. Hydrophobic areas are also shown. The location of the formate anion is as for Fig. 13. H-bonds are indicated by broken lines. The pointer shows the direction of hydride anion transfer.

perturbing the NAD nicotinamide as its OD1 atom is 3·1 Å from C5N. From the side opposite to the nicotinamide the position of the formate anion is fixed by an H-bond to Ile122 N. This interaction is impossible in aFDH as the wall comprising the Ile122-Gly123 pair as well as several neighbouring residues advance about 1·8 Å towards the substrate binding site on transition from aFDH to hFDH (Fig. 15). Arg284 is about 0·8 Å closer to the active site in hFDH compared to aFDH. The conformation of its side-chain is stabilized by H-bonds with Ala283 O and Gly123 O. In aFDH Arg284 is poorly ordered and interacts with Ala283 O and Asp125 OD1, OD2. The Arg284 side-chain becomes properly oriented by making a new contact with Gly123 O, only possible in hFDH. The fact that Arg284 acquires a new co-ordination suggests that it is important for catalysis. Its role is to bind formate (in the first location) or to stabilize the negatively charged hydride in the transition state (in the second location) and probably also to interact with the nicotinamide ring perturbing its electronic state: the distance Arg284 NH1 to the nearest C2N is 3·3 Å. Regardless of the fine details of formate location the role of the "positively" charged cluster is mostly to bind the formate anion and to fix it in the proper position relative to NAD.

There are several hydrophobic clusters in the

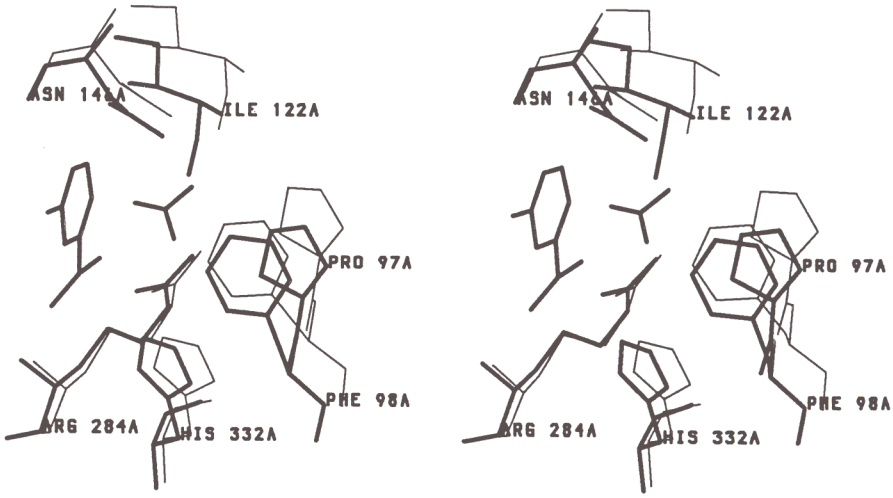


Figure 15. Stereo view showing shifts of the active site residues. hFDH with nicotinamide and formate in its first location are in bold, aFDH is drawn as a thin line.

active site (Fig. 13(d)). The first is composed of residues Ile202 and Val150 pointing directly towards one side of the NAD pyridine ring and providing a hydrophobic environment for it. The position of these residues in hFDH differs by less than 0·4 Å from those in aFDH. The second cluster comprises Pro97 and Phe98. Their side-chains in hFDH move more than 1 Å and point to the formate binding site (Fig. 15). The H-bond Pro97 O to Gln313 NE2 (3·0 Å) occurs in hFDH only: in aFDH the corresponding distance is 4·5 Å. The third hydrophobic region is the Ile122 side-chain hanging over the edge of the pyridine ring. The fourth hydrophobic area is made by the Val309 side-chain and supports the proper conformation of the NAD carboxamide.

His332 is another key residue of the FDH active centre. Its side-chain moves by 0·6 Å closer to C4N in hFDH compared to aFDH. The role of this histidine is controversial. His332 NE2 atom could form an H-bond with formate in the second location provided the whole side-chain is moved or rotated and the histidine acts as one of the substrate ligands. However, in both aFDH and hFDH His332 NE2 does not form any H-bond. Moreover its ND1 atom forms a well-defined (2·8 Å) H-bond with Gln313 NE2. This means that ND1 cannot be protonated otherwise it would have too close contacts with the hydrogen atoms of Gln313 NE2: the distance from the hypothetical hydrogen on His332 ND1 to one of hydrogen atoms at Gln313 NE2 is 1·0 Å. The identities of Gln313 NE2 and OE1 are certain in the crystal structure firstly due to the well-defined H-bonds Gln313 OE1 to Trp310 N (2·8 Å), Gln313 NE2 to Pro97 O (3·0 Å) and secondly because of consistency in the temperature factors for the Gln313 side-chain atoms. Therefore, in hFDH only His332 has a hydrogen at the NE2 position and is uncharged. This is particularly significant if the *cis* conformations of Pro312 and Pro314 are taken into account. These are the only *cis* proline residues in both models. They lie on either side of the Gln313 residue, are located in the irregular loop and probably their only function is to provide Gln313 with the proper orientation to make it impossible for His332 to be positively charged. A possible role of His332 could be to perturb the NAD nicotinamide unit through interaction with its carboxamide as the distance between His332 NE2 and NAD C7N is 3·0 Å. On the other hand the His332 side-chain is located near the Phe98 side-chain and together with Pro97 forms a hydrophobic zone. The proposed hydrophobicity of His332 is based on the fact that its imidazole moiety cannot be charged.

(e) *Alignment with* D*-enzymes*

Recently a sequence alignment of D-specific dehydrogenases was carried out for D-lactate, D-2-hydroxyisocaproate, glycerate, D-3-phosphoglycerate and erythronate 4-phosphate dehydrogenases (Kochhar *et al.*, 1992). All these catalyse dehydrogenation of D-hydroxy acids. It was concluded that these proteins belong to the D-specific 2-hydroxy acid dehydrogenase family. This family was extended by addition of two more D-lactate dehydrogenases, vancomycin resistance protein and FDH (Vinals *et al.*, 1993). Homologous regions were detected in the β-sheet in the FDH coenzyme binding domain and in the interdomain contact region. Several residues (in the FDH sequence Asp249, Lys274, Asn281, Arg284, Asp289, Leu297 and His332) were found to be fully conserved. Of these only Arg284 and His332 are part of the FDH active site as defined from the X-ray structure. Leu297, together with other hydrophobic residues, is part of the hydrophobic core, near to the NAD adenine binding site, and its conservation may be accidental. Asn281 does not directly participate in the active site in spite of the fact that its neighbours do. Asp249, Lys274 and Asp289 in FDH are located 20 to 25 Å away from the nearest NAD atom.

Table 11

Alignment of the residues participating in the FDH active site with NAD-dependent formate dehydrogenases and some D*-specific enzymes.*

	Catalytic domain				Interdomain region		Coenzyme binding domain							Interdomain region		
Residue number	97	98	122	123	146	150	202	282	283	284	308	309	313	332	334	335
FDH	P	F	I	G	N	V	I	T	A	R	D	V	Q	H	S	G
YFDH	P	F	V	G	N	V	I	T	A	R	D	V	Q	H	S	G
PFDH	P	F	I	G	N	V	I	N	A	R	D	V	Q	H	S	G
DLDH1	V	V	V	G	S	I	I	V	S	R	D	V	E	H	A	F
DLDH2	A	D	V	G	S	I	I	F	A	R	V	T	E	H	A	F
DHICDH	I	N	V	G	S	I	I	T	A	R	D	T	E	H	A	Y
VANDH	A	L	I	T	S	V	I	T	G	R	D	V	E	H	A	Y
GLYDH	V	I	V	G	L	T	I	C	S	R	D	V	E	H	A	S
PGDH	H	F	I	G	N	V	I	A	S	R	D	V	E	H	G	G
ERYDH	L	A	A	G	N	V	V	A	C	R	D	V	E	H	A	G

D-Specific enzymes are aligned as suggested by Vinals *et al.* (1993). Numbers for the FDH sequence are given. Abbreviations are: YFDH, yeast formate dehydrogenase from *Hansenula polimorpha*, PFDH, potato formate dehydrogenase, DLDH1, D-lactate dehydrogenase from *Lactobacillus delbrueckii*, DLDH2, D-lactate dehydrogenase from *Lactobacillus plantarium*, DHICDH, D-2-hydroxyisocaproate dehydrogenase from *Lactobacillus casei*, VANDH, vancomycin resistance protein from *Enterococcus faecium*, GLYDH, glycerate dehydrogenase from *Cucumis saivus*, PGDH, D-3-phosphoglycerate dehydrogenase from *E. coli*, ERYDH, erythronate 4-phosphate dehydrogenase from *E. coli*.

The similarity between FDH and the D-enzymes suggest that these proteins may have similar folds. Here only the residues in the FDH active site are compared (Table 11). They are numbered as in the FDH sequence. NAD-dependent formate dehydrogenases from yeast (Hollenberg & Janowicz, 1989) and potato (Colas des Francs, personal communication), which have about 45% identity with the FDH sequence, have been added to the alignment and with two exceptions (Ile122 to Val in yeast and Thr282 to Asn in potato enzymes) have the same residues as those in the FDH active site.

(i) *Residues interacting with the NAD carboxamide*

Asp308 is essentially conserved and is probably responsible for proper NAD binding in all D-enzymes. Lactate dehydrogenase from *Lactobacillus plantarium* (DLDH2) is an exception: there is valine at this position. Thr282 only interacts with NAD through its carbonyl O and hence its side-chain is not conserved. It can be replaced by other residues such as cysteine, alanine, asparagine or even phenylalanine as in DLDH2. Residue 335, which interacts through its main-chain atoms, also has a non-conserved side-chain, which can be either serine, phenylalanine, tyrosine or glycine. Surprisingly Ser334, whose OG makes an H-bond to the NAD carboxamide, is present in formate dehydrogenases only. However, at this position the side-chain is always small (alanine or glycine). Val309, supporting proper orientation of the carboxamide unit by the interaction CG1 to NAD N7N in FDH, is nearly conserved in D-enzymes. It is replaced by threonine in D-2-hydroxyisocaproate and lactate (*L. plantarium*) dehydrogenases. However, threonine is similar to valine and its CG2 atom is spatially equivalent to CG1 in valine.

(ii) *Residues binding the substrate*

Formate is the smallest substrate among enzymes of the "D-specific 2-hydroxy acid dehydrogenase" family. Reaction with formate differs from that with 2-hydroxy acids in two respects: the hydride anion leaves the first carbon atom (counting from the carboxyl group) compared to 2-hydroxy acids where hydride leaves the second carbon. Secondly the formate molecule is not chiral. This suggests that active sites of enzymes acting on 2-hydroxy acids should differ from the FDH active site if residues binding the substrate are considered.

Arg284 is fully conserved in all aligned sequences. This supports the conclusion that this is one of the most important residues for catalysis. In this respect the first location of formate bound to Arg284 is again preferred. The orientation of the Arg284 side-chain is maintained by Ala283 O and Gly123 O. The side-chain of residue 283 varies and is small, either alanine, serine, cysteine or glycine. Gly123 is nearly completely conserved. In FDH it has left-handed helical dihedral angles, so the presence of any side-chain would cause steric hindrance with neighbouring main-chain atoms and the Arg284 side-chain. However, it is replaced by threonine in vancomycin. Ile122, participating in substrate binding by its main-chain atoms and providing a hydrophobic area by its side-chain, is conserved in terms of hydrophobicity and is in several proteins replaced by valine or alanine. The alignment of Ile122 and Gly123 should be treated with caution as they belong to the catalytic domain where the homology is low. Asn146 ND2 provides an H-bond to substrate in FDH. In D-enzymes it is not conserved and is either asparagine, serine or leucine. However, if the first location of formate between Arg284 and Asn146 side-chains is assumed,

the Asn146 does not need to be conserved in D-enzymes because of the stereochemical and geometrical differences between dehydrogenation of the formate and 2-hydroxy acids.

(iii) *Hydrophobic cores*

Besides Ile122 and Val309 there are two extensive hydrophobic areas in the FDH active site. Val150 and Ile202 point to the nicotinamide unit. Both these residues are conserved in D-enzymes in terms of hydrophobicity. Several proteins contain isoleucine or threonine instead of Val150. Ile202 is essentially conserved and is changed to valine only in erythronate 4-phosphate dehydrogenase. The second cluster facing the substrate, Pro97-Phe98, is specific to formate dehydrogenases and appears not to be conserved at all in D-enzymes. These two residues belong to the catalytic domain which shows low homology.

(iv) *Histidine residue*

His332 in FDH is fully conserved in all D-enzymes. These enzymes need a general acid–base for catalysis and histidine is the most probable candidate. An Asp-His pair linked by an H-bond and functioning as a "charge relay system" was found in L-specific dehydrogenases (Birktoft & Banaszak, 1983). On the basis of sequence alignment a similar Asp-His pair was suggested for D-specific enzymes (Kochhar *et al.*, 1992) and in FDH (Asp308-His332: Vinals *et al.*, 1993). However, according to the three-dimensional structure His332 in FDH is not involved in such a pair. His332 makes a single H-bond ND1 to Gln313 NE2 and is uncharged. All D-enzymes have glutamate instead of glutamine at the equivalent position. Assuming similarity in three-dimensional structure, this glutamate forms an H-bond with histidine ND1, as in FDH, the histidine in the D-enzymes has a hydrogen at the ND1 position and can be positively charged with hydrogen atoms both at ND1 and NE2. Thus, D-specific enzymes seem to have the Glu-His "charge relay system" instead of Asp-His. We suggest the presence of histidine in the FDH active site supports an evolutionary relation to D-specific dehydrogenases. Its ability to accept hydrogen at ND1 and the resulting possibility of switching between charged and uncharged forms, realized in various catalytic mechanisms, is locked in FDH by the Glu313 to Gln313 replacement.

The results derived from sequence alignment should be treated carefully because even identity at a position in the sequence does not prove superposition in three dimensions. Additional disagreements may arise from imperfect alignment of sequences. However, such alignment may be used to support hypotheses about the role of a particular residue. Thus, the suggested role of all the hydrophilic residues participating in the FDH active site agrees with the alignment. The presence of the hydrophobic cluster, Val150 and Ile202, as well as the hydrophobic side-chains for Ile122 and Val309 is also in good agreement. The cluster made up by Pro97 and Phe98 from the catalytic domain in FDH is not conserved in the alignment. The alignment does not extend to the catalytic domain.

The formate anion, the substrate in FDH catalysis, is not chiral and could well be considered as an example of either L or D-substrates. However, the conservation found for most of the residues in the FDH active site supports the idea that FDH belongs to the "D-specific 2-hydroxy acid dehydrogenase" family.

(f) *Catalytic mechanism*

In the transition state of the reaction catalysed by FDH a hydride anion leaves the formate and attacks the electrophilic C4N of the positively charged nicotinamide moiety of NAD. As a result two neutral species, CO_2 and NADH, are produced while NAD C4N changes its hybridization from sp_2 to sp_3 and the nicotinamide moiety becomes uncharged. Knowledge of the three-dimensional crystal structures of the two forms of FDH allows the identification of the residues participating in substrate binding and the active site. Accurate determination of the atomic positions pinpoints important H-bonds and other interatomic interactions. This allows us to propose a mechanism for FDH catalysis.

Two possible mechanisms are discussed arising from the two possible locations of formate in the active centre. In spite of substantial differences there are some similar features for both alternative mechanisms of FDH action. Whatever the mode of formate binding the mechanism of H-transfer in FDH catalysed reaction appears to be governed by electrostatic effects. The positively charged NAD nicotinamide unit is properly oriented in the active centre by multiple interactions with a "negatively" charged cluster and hydrophobic side-chains, whereas formate is bound and oriented by multiple H-bonds with a "positively" charged cluster. The most important residues for formate binding are Arg284 and Asn146 in the first variant of the mechanism and His332 and Asn146 in the second, corresponding to the two locations of formate. Main-chain atoms from a number of other residues support the H-bond scheme in the active site. Hydrophobic areas in the vicinity of the substrate and nicotinamide play an important role in catalysis. They are located on different sides of the active site and surround the substrate and the nicotinamide as shown schematically in Figure 16.

An important factor for catalysis by all NAD-dependent dehydrogenases including FDH should be enhancement of the electrophilic properties of the C4N atom of the nicotinamide moiety of the co-enzyme. It might be achieved through sufficient polarization of the NAD carboxamide group and perturbation of its ground state due to the twist of the carboxamide with respect to the pyridine plane.

According to energy calculations the most stable

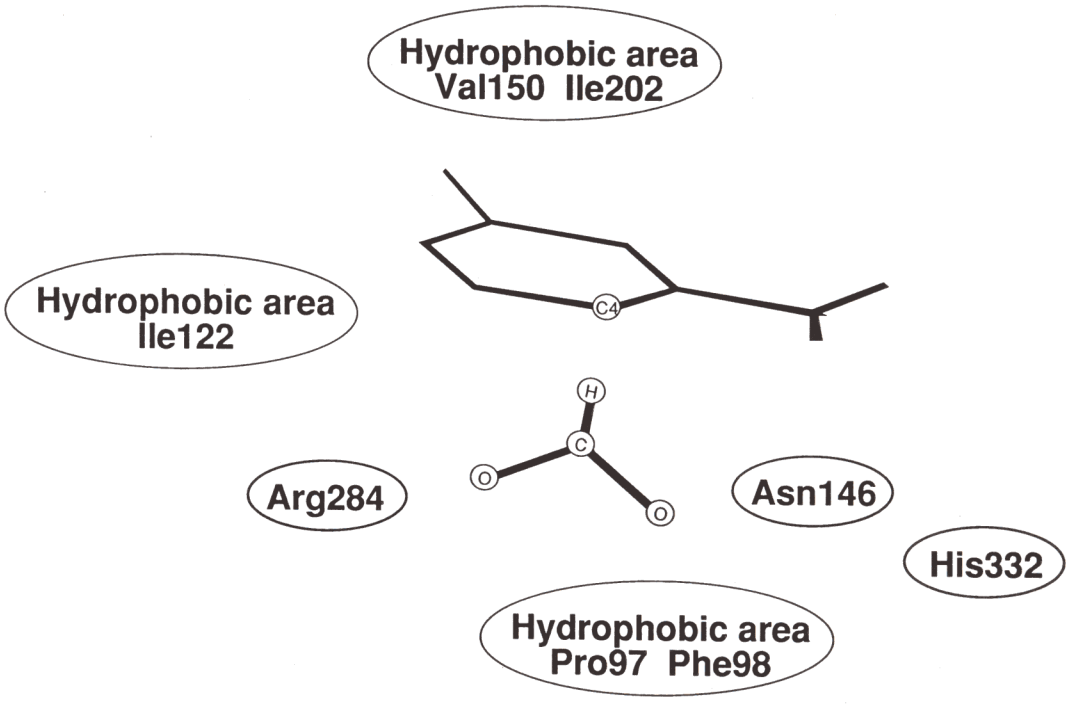


Figure 16. Schematic representation of the FDH active site. The negatively charged formate anion is properly oriented *versus* the positively charged nicotinamide unit by the "positively" charged cluster. Hydrophobic areas facilitate the hydride transfer.

conformation for non-bound NAD is when the carboxamide group is in the *trans* conformation (a torsion angle is defined by atoms O7N-C7N-C3N-C4N) and is coplanar with the pyridine ring (Li & Goldstein, 1992). The *cis* conformation is less preferable and is about 2 kcal/mol less stable for the torsion angle being between $-30°$ to $+30°$. However, the *cis* conformation supports a polarization of the pyridine ring and a partial positive charge on the C4N atom. In FDH the carboxamide group is in the *cis* conformation and makes a torsion angle of $+25°$ with the pyridine plane. This arises from H-bonding and non-bonding interactions between carboxamide and neighbouring atoms forcing the carboxamide plane to be non-planar with the pyridine and causing a perturbation of the nicotinamide electronic state. Such an activation is common within dehydrogenases. Most of those have NAD(P)/NAD(P)H bound with carboxamide in the *cis* conformation where the torsion angle varies from $-27°$ to $+34°$ (Li & Goldstein, 1992).

The central dogma of catalysis assumes that the transition state is bound better than reactants. The catalytic effect will also be achieved if the enzyme perturbs the ground state of reactants and thus lowers the activation barrier of the process. Several interactions resulting in the stabilization of the hypothetical transition state of either forward or reverse FDH reaction are realized for both mechanisms: interaction of the positively charged Arg284 side-chain with the negatively charged hydride ion transferred in the transition state, polarization of NAD C4N through multiple interactions with a "negatively" charged zone and co-ordination of the carbon atom of formate/CO_2 by Ile122 O resulting in polarization of the formate C–H bond. A number of possible destabilizing interactions can be equally important: destabilization of the formate ground state due to interaction of its oxygen atoms with the hydrophobic environment (Pro97-Phe98) and Ile122 O and destabilization of the nicotinamide ground state due to interaction with His332 (mechanism I) or Arg284 (mechanism II) side-chains. Thus, the main difference between the two hypothetical mechanisms of FDH appears to be the different roles ascribed to Arg284 and His332. In either case these two play an essential role in maintaining the general architecture of the enzyme active centre.

Hydrophobic interactions are supposed to play a crucial role in FDH catalysis. Transformation of formate to CO_2 and NAD to NADH can be regarded as a combination of positive and negative charges resulting in two neutral compounds. Thus, if the charges are placed at a proper distance, the hydrophobic environment will support their neutralization. Three walls comprising Pro97/Phe98, Ile122 and Val150/Ile202 are pressed over the active site as a result of the conformational change in the course of the reaction. The Pro-Phe pair may perturb the formate ground state and provide a more comfortable environment for neutral CO_2, while the Val150 and Ile202 side-chains face the other side of the pyridine ring and may be responsible for the nicotinamide in NADH acquiring a periplanar or antiperiplanar boat conformation (Almarsson *et al.*, 1992).

The proposed mechanisms cannot be put forward with certainty until high resolution structures of

FDH with other ligands (formate, NADH) are available. Site directed mutagenesis of residues supposed to be catalytically important on the basis of the aFDH and hFDH crystal structures, followed by biochemical studies or structure determination of mutants, will also provide important information. These studies are now in progress.

This work was supported in part by an EMBL postdoctoral fellowship to V.S.L. and a grant from Russian Ministry of Science, High School and Technology. We thank Dr Catherine Colas des Francs from Paris University for providing the sequence of potato formate dehydrogenase prior to its publication. The co-ordinate and X-ray data for both aFDH and hFDH have been submitted to the Protein Data Bank (Bernstein *et al.*, 1977) under accession numbers 1NAC for the aFDH co-ordinates, 1NAC-SF for the aFDH X-ray data, 1NAD for the hFDH co-ordinates and 1NAD-SF for the hFDH X-ray data.

References

Abad-Zapatero, C., Griffith, J. P., Sussman, J. L. & Rossmann, M. G. (1987). Refined crystal structure of dogfish M_4 apo-lactate dehydrogenase. *J. Mol. Biol.* **198**, 445–467.

Adams, M. J., Ford, G. C., Koekoek, R., Lentz, P. J., McPherson, A., Rossmann, M. G., Smiley, I. E., Shewitz, R. W. & Wonacott, A. J. (1970). Structure of lactate dehydrogenase at 2·8 Å resolution. *Nature (London)*, **227**, 1098–1103.

Agarwal, R. C. (1978). A new least-squares refinement technique based on the fast Fourier transform algorithm. *Acta Crystallogr. sect. A*, **34**, 791–809.

Almarsson, Ö., Karaman, R. & Bruice, T. C. (1992) Kinetic importance of conformations of nicotinamide adenine dinucleotide in the reaction of dehydrogenase enzymes. *J. Amer. Chem. Soc.* **114**, 8702–8704.

Archimedes (230 BC). Quoted in *Pappus Alexandr., Collectio* (1878), (Hultsch, ed.), **8**, prop. 10, § 11.

Baker, E. N. & Dodson, E. J. (1980). Crystallographic refinement of the structure of actinidin at 1·7 Å resolution by fast Fourier least-squares methods. *Acta Crystallogr. sect. A*, **36**, 559–572.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.

Birktoft, J. J. & Banaszak, L. J. (1983). The presence of a histidine-aspartic acid pair in the active site of 2-hydroxyacid dehydrogenases. *J. Biol. Chem.* **258**, 472–482.

Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C. & Kraut, J. (1982). Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1·7 Å resolution. I. General features and binding of methotrexate. *J. Biol. Chem.* **257**, 13650–13662.

Bricogne, G. (1974). Geometric sources of redundancy in intensity data and their use for phase determination. *Acta Crystallogr. sect. A*, **30**, 395–405.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science*, **235**, 458–460.

Bystroff, C. & Kraut, J. (1991). Crystal structure of unliganded *Escherichia coli* dihydrofolate reductase. Ligand-induced conformational changes and co-operativity in binding. *Biochemistry*, **30**, 2227–2239.

CCP4 (1979). *The SERC (UK) Collaborative Computing Project No. 4*: a suite of programs for protein crystallography, distributed from Daresbury laboratory. Warrington WA4 4AD, UK.

Cygler, M. & Anderson, W. F. (1988). Application of the molecular replacement method to multidomain proteins. 1. Determination of the orientation of an immunoglobulin Fab fragment. *Acta Crystallogr. sect. A*, **44**, 38–45.

Dauter, Z., Betzel, C., Genov, N., Pipon, N. & Wilson, K. S. (1991). Complex between the subtilisin from a mesophilic bacterium and the leech inhibitor eglin-C. *Acta Crystallogr. sect. B*, **47**, 707–730.

Davies, D. R. & Segal, D. M. (1971). Protein crystallization: micro techniques involving vapour diffusion. In *Methods Enzymol.* (Jakoby, W. B., ed.), vol. 22, pp. 266–269, Academic Press, New York and London.

Dunn, C. R., Wilks, H. M., Halsall, D. J., Atkinson, T., Clarke, A. R., Muirhead, H. & Holbrook, J. J. (1991). Design and synthesis of new enzymes based on the lactate dehydrogenase framework. *Phil. Trans. Roy. Soc. Ser. B*, **332**, 177–184.

Egorov, A. M., Avilova, T. V., Dikov, M. M., Popov, V. O., Rodionov, Yu. V. & Berezin, I. V. (1979). NAD-dependent formate dehydrogenase from methylotrophic *Bacterium*, strain 1. *Eur. J. Biochem.* **99**, 569–576.

Eklund, H. & Bränden, C.-I. (1987). Crystal structure, coenzyme conformations and protein interactions. In *Pyridine Nucleotide Coenzymes* (Dolphin, D. N. Y., ed.) pp. 51–98, Wiley, New York.

Grau, U. M., Trommer, W. E. & Rossmann M. G. (1980). Structure of the active ternary complex of pig heart lactate dehydrogenase with *S*-lac-NAD at 2·7 Å resolution. *J. Mol. Biol.* **151**, 289–307.

Hermes, J. D., Morrical, S. W., O'Leary, M. H. & Cleland, W. W. (1984). Variation of transition-state structures as a function of the nucleotide in reactions catalyzed by dehydrogenases. 2. Formate dehydrogenase. *Biochemistry*, **23**, 5479–5488.

Hodges, C. T., Wood, D. C. & Harrison, J. H. (1978). Investigation of the antico-operative binding of NADPH to porcine heart mitochondrial malate dehydrogenase. *J. Biol. Chem.* **253**, 4854–4864.

Hollenberg, C. P. & Janowicz, Z. (1989). European patent. Publication number 0 299 108. Bulletin 89/03. Publication date January 18, 1989.

Iwata, S. & Otha, T. (1993). Molecular basis of allosteric activation of bacterial L-lactate dehydrogenase. *J. Mol. Biol.* **230**, 21–27.

Jones, T. A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–272.

Karplus, P. A. & Schulz, G. E. (1987). Refined structure of glutathione reductase at 1·54 Å resolution. *J. Mol. Biol.* **195**, 701–729.

Kochhar, S., Hunziker, P. E., Leong-Morgenthaler, P. & Hottinger, H. (1992). Evolutionary relationship of NAD^+-dependent D-lactate dehydrogenase: comparison of primary structure of 2-hydroxy acid dehydrogenases. *Biochem. Biophys. Res. Commun.* **184**, 60–66.

Konnert, J. H. & Hendrickson, W. A. (1980). A restrained-parameter thermal-factor refinement procedure. *Acta Crystallogr. sect. A*, **36**, 344–350.

Lamzin, V. S. & Wilson, K. S. (1993). Automated refine-

ment of protein models. *Acta Crystallogr. sect. D*, **49**, 129–147.

Lamzin, V. S., Asadchikov, V. E., Popov, V. O., Egorov, A. M. & Berezin I. V. (1986). Conformational changes of the bacterial formate dehydrogenase on binding of the cofactors and their analogues. *Dokl. Akad. Nauk U.S.S.R.* **291**, 1011–1014.

Lamzin, V. S., Aleshin, A. E., Strokopytov, B. V., Yukhnevich, M. G., Popov, V. O., Harutyunyan, E. H. & Wilson, K. S. (1992). Crystal structure of NAD-dependent formate dehydrogenase. *Eur. J. Biochem.* **206**, 441–452.

Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, **26**.

Li, H. & Goldstein, B. M. (1992). Carboxamide group conformation in the nicotinamide and thiazole-4-carboxamide rings: implications for enzyme binding. *J. Med. Chem.* **35**, 3560–3567.

Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). Stereochemical quality of protein structure co-ordinates. *Proteins: Struct. Funct. Genet.* **12**, 345–364.

Oefner, C., D'Arcy, A. & Winkler, F. K. (1988). Crystal structure of human dihydrofolate reductase complexed with folate. *Eur. J. Biochem.* **174**, 377–385.

Popov, V. O., Rodionov, Yu. V., Egorov, A. M. & Berezin, I. V. (1978). Formate dehydrogenase from Bacterium sp. 1: number of active sites and coenzyme dissociation constants. *Dokl. Akad. Nauk U.S.S.R.* **239**, 1482–1485.

Popov, V. O., Shumilin, I. A., Ustinnikova, T. B., Lamzin, V. S. & Egorov, Ts. A. (1990). NAD-dependent formate dehydrogenase from methylotrophic bacterium *Pseudomonas sp.* 101. I. Amino acid sequence. *Bioorg. Khimia (U.S.S.R.)* **16**, 324–335.

Ptitsyn, O. B., Sinev, M. A., Razgulaev, O. I., Pavlov, M. Yu. & Timchenko, A. A. (1988). Interdomain mobility of enzymes and its functional role. *J. Mol. Catal.* **47**, 289–295.

Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. sect. A*, **42**, 140–149.

Rossmann, M. G., Liljas, A., Bränden, C.-I. & Banaszak, L. J. (1975). Evolutionary and structural relationships among dehydrogenases. In *The Enzymes* (Boyer, P. D. ed.), 3rd. edit., vol. 11A, pp. 61–102, Academic Press, New York, London.

Scrutton, N. S., Berry, A. & Perham, R. N. (1990). Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature (London)*, **343**, 38–43.

Skarzynski, T., Moody, P. C. E. & Wonacott, A. J. (1987). Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1·8 Å resolution. *J. Mol. Biol.* **193**, 171–187.

Sussman, J. L., Holbrook, S. R., Church, G. M. & Kim, S.-H. (1977). A structure factor least-squares refinement procedure for macromolecular structures using constrained and restrained parameters. *Acta Crystallogr. sect. A*, **33**, 800–804.

Tishkov, V. I., Galkin, A. G. & Egorov, A. M. (1991). NAD-dependent formate dehydrogenase from the methylotrophic bacterium *Pseudomonas* sp. 101: cloning, expression and studying of the gene structure. *Dokl. Akad. Nauk U.S.S.R.* **317**, 745–748.

Ustinnikova, T. B., Popov, V. O. & Egorov, Ts. A. (1988) Structural studies of the NAD-dependent formate dehydrogenase from methylotrophic bacterium. Localisation of the essential cysteine. *Bioorg. Khimia (U.S.S.R.)*, **14**, 905–909.

Vinals, K., Depiereux, E. & Feytmans, E. (1993). Prediction of structurally conserved regions of D-specific hydroxy acid dehydrogenases by multiple alignment with formate dehydrogenase. *Biochem. Biophys. Res. Commun.* **192**, 182–188.

Vrielink, A., Lloyd, L. F. & Blow, D. M. (1991). Crystal structure of cholesterol oxidase from *Brevibacterium sterolicum* refined at 1·8 Å resolution. *J. Mol. Biol.* **219**, 533–554.

Wierenga, R. K., de Maeyer, M. C. H. & Hol, W. G. J. (1985). Interaction of pyrophosphate moieties with α-helixes in dinucleotide binding proteins. *Biochemistry*, **24**, 1346–1357.

Wilmot, C. M. & Thornton, J. M. (1990). β-Turns and their distortions: a proposed new nomenclature. *Protein Eng.* **3**, 479–493.

Wilson, A. J. C. (1942). Determination of absolute from relative X-ray data intensities. *Nature (London)*, **150**, 151–152.

Edited by W. Hendrickson

(Received 15 June 1993; accepted 26 October 1993)